

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:51:16 ; Search time 10.6667 Seconds
(without alignments)
19.309 Million cell updates/sec

Title: US-10-006-069a-55

Perfect score: 35

Sequence: 1 DRSNLTR 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	89	3	US-08-793-408-18
2	35	100.0	89	3	US-09-139-762A-18
3	31	88.6	47	4	US-09-187-789-43
4	31	88.6	47	4	US-09-139-600-38
5	31	88.6	278	3	US-08-522-813-4
6	31	88.6	293	1	US-08-446-925-5
7	31	88.6	293	1	US-08-146-331-5
8	31	88.6	293	2	US-08-896-885-5
9	31	88.6	293	4	US-09-375-256-5
10	31	88.6	293	4	US-09-561-756-21
11	31	88.6	293	4	US-09-227-721-21
12	31	88.6	293	4	US-08-983-502-31
13	31	88.6	293	4	US-09-376-156-5
14	31	88.6	293	5	PCT-US96-10521-31
15	31	88.6	300	4	US-09-561-756-36
16	31	88.6	300	4	US-09-227-721-36
17	30	85.7	10	3	US-09-139-762A-72
18	30	85.7	10	3	US-09-139-762A-75
19	30	85.7	10	3	US-09-139-762A-93
20	30	85.7	13	4	US-08-842-306B-30
21	30	85.7	13	4	US-08-838-973B-30
22	30	85.7	13	4	US-08-771-212A-26
23	30	85.7	33	3	US-08-793-408-12
24	30	85.7	33	3	US-09-139-762A-12
25	30	85.7	377	1	US-08-188-277B-4
26	30	85.7	377	2	US-08-429-964-78
27	30	85.7	451	4	US-09-357-251-35

28	30	85.7	1523	4	US-09-182-024A-2	Sequence 2, Appli
29	29	82.9	10	3	US-09-139-762A-30	Sequence 30, Appl
30	29	82.9	10	3	US-09-139-762A-62	Sequence 62, Appl
31	29	82.9	10	3	US-09-139-762A-70	Sequence 70, Appl
32	29	82.9	13	4	US-08-842-306B-29	Sequence 29, Appl
33	29	82.9	13	4	US-08-838-973B-29	Sequence 29, Appl
34	29	82.9	13	4	US-08-771-212A-25	Sequence 25, Appl
35	29	82.9	276	4	US-09-068-195-11	Sequence 11, Appl
36	29	82.9	276	4	US-09-068-195-23	Sequence 23, Appl
37	29	82.9	377	1	US-08-188-277B-2	Sequence 2, Appli
38	29	82.9	377	2	US-08-429-964-80	Sequence 80, Appl
39	28	80.0	822	2	US-08-222-617A-7	Sequence 7, Appli
40	28	80.0	3666	2	US-08-222-617A-12	Sequence 12, Appl
41	28	80.0	3727	2	US-08-222-617A-27	Sequence 27, Appl
42	28	80.0	3778	2	US-08-222-617A-2	Sequence 2, Appli
43	27	77.1	375	1	US-08-303-238-2	Sequence 2, Appli
44	27	77.1	375	4	US-08-458-834-2	Sequence 1, Appli
45	27	77.1	376	1	US-08-303-238-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-793-408-18
; Sequence 18, Application US/08793408
; Patent No. 6007988
; GENERAL INFORMATION:
; APPLICANT: Choo, Yen
; APPLICANT: Klug, Aaron
; APPLICANT: Sanchez Garcia, Isidro
; TITLE OF INVENTION: Improvements in or Relating to
; TITLE OF INVENTION: Binding Proteins for Recognition of DNA
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,408
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/01949
; FILING DATE: 17-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9514698.1
; FILING DATE: 18-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9422534.9
; FILING DATE: 08-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9416880.4
; FILING DATE: 20-AUG-1994
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-793-408-18
Query Match 100.0%; Score 35; DB 3; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.77;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 7
|||||||
Db 74 DRSNLT 80

RESULT 2

US-09-139-762A-18
; Sequence 18, Application US/09139762A
; Patent No. 6013453
; GENERAL INFORMATION:
; APPLICANT: Choo, Yen
; APPLICANT: Klug, Aaron
; APPLICANT: Sanchez Garcia, Isidro
; TITLE OF INVENTION: Improvements in or Relating to
; TITLE OF INVENTION: Binding Proteins for Recognition of DNA
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/139,762A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/793,408
; FILING DATE: 02-JUN-1997
; APPLICATION NUMBER: PCT/GB95/01949
; FILING DATE: 17-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9514698.1
; FILING DATE: 18-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9422534.9
; FILING DATE: 08-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9416880.4
; FILING DATE: 20-AUG-1994
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-139-762A-18

Query Match 100.0%; Score 35; DB 3; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.77; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 7
|||||||
Db 74 DRSNLT 80

RESULT 3

US-09-187-789-43
; Sequence 43, Application US/09187789
; Patent No. 6340740
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; APPLICANT: Fernandez-Alnemri, Teresa

; TITLE OF INVENTION: CASPASE-14, AN APOTOTIC PROTEASE, NUCLEIC ACID ENCODING
; FILE REFERENCE: 480140.434C1
; CURRENT APPLICATION NUMBER: US/09/187,789
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-187-789-43

Query Match 88.6%; Score 31; DB 4; Length 47;
Best Local Similarity 85.7%; Pred. No. 2.9;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLT 7
|||||||
Db 12 DRDNLT 18

RESULT 4

US-09-139-600-38
; Sequence 38, Application US/09139600
; Patent No. 6432628
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; APPLICANT: Fernandez-Alnemri, Teresa
; TITLE OF INVENTION: CASPASE-14, AN APOTOTIC PROTEASE, NUCLEIC ACID ENCODING
; FILE REFERENCE: 480140.434
; CURRENT APPLICATION NUMBER: US/09/139,600
; CURRENT FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-139-600-38

Query Match 88.6%; Score 31; DB 4; Length 47;
Best Local Similarity 85.7%; Pred. No. 2.9;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLT 7
|||||||
Db 12 DRDNLT 18

RESULT 5

US-08-522-813-4
; Sequence 4, Application US/08522813
; Patent No. 6033848
; GENERAL INFORMATION:
; APPLICANT: Braxton, Scott M
; APPLICANT: Diep, Dinh
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: HUMAN ICE HOMOLOG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3330 Hillview Avenue
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/522,813
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Luther, Barbara J.
;; REGISTRATION NUMBER: 33954
;; REFERENCE/DOCKET NUMBER: PF-0045P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-855-0555
;; TELEFAX: 415-852-0195
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 278 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-522-813-4

Query Match 88.6%; Score 31; DB 3; Length 278;
Best Local Similarity 85.7%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 DRSNLTR 7
Db 55 DRDLNTR 61

RESULT 6
US-08-446-925-5
; Sequence 5, Application US/08446925
; Patent No. 5672500
; GENERAL INFORMATION:
; APPLICANT: Litwack, Gerald
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: Fernandez-Alnemri, Teresa
; TITLE OF INVENTION: Mch2, AN APOPTOTIC CYSTEINE
; TITLE OF INVENTION: PROTEASE,
; TITLE OF INVENTION: AND COMPOSITIONS FOR MAKING AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESSEE: No. 5672500ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,925
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1508
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 293 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-446-925-5

Query Match 88.6%; Score 31; DB 1; Length 293;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
Db 70 DRDLNTR 76

RESULT 7
US-09-146-331-5
; Sequence 5, Application US/09146331
; Patent No. 5958720
; GENERAL INFORMATION:
; APPLICANT: Litwack, Gerald
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: Fernandez-Alnemri, Teresa
; TITLE OF INVENTION: Mch2, AN APOPTOTIC CYSTEINE
; TITLE OF INVENTION: PROTEASE, AND COMPOSITIONS FOR MAKING AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESSEE: No. 5958720ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,331
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/896,885
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1508
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 293 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-146-331-5

Query Match 88.6%; Score 31; DB 2; Length 293;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
Db 70 DRDLNTR 76

RESULT 8
US-08-896-885-5
; Sequence 5, Application US/08896885
; Patent No. 5985640
; GENERAL INFORMATION:
; APPLICANT: Litwack, Gerald
; APPLICANT: Alnemri, Emad S.

APPLICANT: Fernandez-Alnemri, Teresa
; TITLE OF INVENTION: Mch2, AN APOPTOTIC CYSTEINE
; TITLE OF INVENTION: PROTEASE AND COMPOSITIONS FOR MAKING AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESSEE: No. 5985640ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,885
; FILING DATE: 18-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,925
; FILING DATE: 18-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1508
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 293 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-896-885-5

Query Match 88.6%; Score 31; DB 2; Length 293;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLT 7
|| |||||
Db 70 DRDNLT 76

RESULT 9
US-09-375-256-5
; Sequence 5, Application US/09375256
; Patent No. 6359127
; GENERAL INFORMATION:
; APPLICANT: Litwack, Gerald
; Alnemri, Emad S.
; Fernandez-Alnemri, Teresa
; TITLE OF INVENTION: Mch2, AN APOPTOTIC CYSTEINE
; PROTEASE
; AND COMPOSITIONS FOR MAKING AND
; METHODS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESSEE: No. 6359127ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/375,256
; FILING DATE: 16-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,925
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1508
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 293 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-375-256-5

Query Match 88.6%; Score 31; DB 4; Length 293;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLT 7
|| |||||
Db 70 DRDNLT 76

RESULT 10
US-09-561-756-21
; Sequence 21, Application US/09561756
; Patent No. 6376226
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: RECOMBINANT, ACTIVE CASPASES AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: 480140.431
; CURRENT APPLICATION NUMBER: US/09/561,756
; CURRENT FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 09/227,721
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-561-756-21

Query Match 88.6%; Score 31; DB 4; Length 293;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLT 7
|| |||||
Db 70 DRDNLT 76

RESULT 11
US-09-227-721-21
; Sequence 21, Application US/09227721
; Patent No. 6379950
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: RECOMBINANT, ACTIVE CASPASES AND USES
; FILE REFERENCE: 480140.431

; CURRENT APPLICATION NUMBER: US/09/227,721
; CURRENT FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-227-721-21

Query Match: 88.6%; Score 31; DB 4; Length 293;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRSNLR 7
Db 70 DRDLNR 76

RESULT 12
US-08-983-502-31
; Sequence 31, Application US/08983502
; Patent No. 6399327
; GENERAL INFORMATION:
; APPLICANT: David WALLACH
; APPLICANT: Mark P. BOLDIN
; APPLICANT: Tanya M. GONCHAROV
; APPLICANT: Yuri V. GOLTSV
; TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS
; TITLE OF INVENTION: AND OTHER PROTEINS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/983,502
; FILING DATE: 16-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10521
; FILING DATE: 14-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 114,615
; FILING DATE: 16-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 114,986
; FILING DATE: 17-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 115,319
; FILING DATE: 14-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 116,588
; FILING DATE: 27-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117,932
; FILING DATE: 16-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH-19
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 293 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-983-502-31

Query Match: 88.6%; Score 31; DB 4; Length 293;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRSNLR 7
Db 70 DRDLNR 76

RESULT 13
US-09-376-156-5
; Sequence 5, Application US/09376156
; Patent No. 6407215
; GENERAL INFORMATION:
; APPLICANT: Litwack, Gerald
; APPLICANT: Alnemri, Emad S.
; APPLICANT: Fernandez-Alnemri, Teresa
; TITLE OF INVENTION: Mch2, AN APOPTOTIC CYSTEINE
; TITLE OF INVENTION: PROTEASE,
; TITLE OF INVENTION: AND COMPOSITIONS FOR MAKING AND
; TITLE OF INVENTION: METHODS
; TITLE OF INVENTION: OF USING THE SAME
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESSEE: No. 6407215ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/376,156
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,925
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1508
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 293 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-376-156-5

Query Match: 88.6%; Score 31; DB 4; Length 293;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRSNLR 7
Db 70 DRDLNR 76

RESULT 14

PCT-US96-10521-31
; Sequence 31, Application PC/TUS9610521
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS
; AND OTHER PROTEINS
; NUMBER OF SEQUENCES: 34
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10521
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 114,615
; FILING DATE: 16-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 114,986
; FILING DATE: 17-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 115,319
; FILING DATE: 14-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 116,588
; FILING DATE: 27-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117,932
; FILING DATE: 16-APR-1996
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 293 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-10521-31

Query Match 88.6%; Score 31; DB 5; Length 293;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLT 7
DB 70 DRDNLT 76

RESULT 15

US-09-561-756-36
; Sequence 36, Application US/09561756
; Patent No. 6376226
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: RECOMBINANT, ACTIVE CASPASES AND USES
; OF INVENTION: THEREOF
; FILE REFERENCE: 480140.431
; CURRENT APPLICATION NUMBER: US/09/561,756
; CURRENT FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 09/227,721
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Rev-Caspase-6 constructed from human caspase-6
US-09-561-756-36

Query Match 88.6%; Score 31; DB 4; Length 300;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLT 7
DB 191 DRDNLT 197

Search completed: July 15, 2003, 12:56:08
Job time : 11.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:48:06 ; Search time 33.3333 Seconds
(without alignments)
27.983 Million cell updates/sec

Title: US-10-006-069A-68

Perfect score: 36

Sequence: 1 RSDHLSR 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	7	AAE08729	Human KCA4 protein
2	36	100.0	7	AAE08729	Human KCA4 protein
3	36	100.0	7	AAE08729	Amino acid sequenc
4	36	100.0	7	AAE08729	F3 zinc finger for
5	36	100.0	7	AAE08729	Human VEGF-targete
6	36	100.0	7	AAE08729	Human VEGF-targete
7	36	100.0	7	AAE08729	Human VEGF-targete
8	36	100.0	7	AAE08729	Human VEGF-targete
9	36	100.0	7	AAE08729	Human VEGF-targete
10	36	100.0	7	AAE08729	Human VEGF-targete

11	36	100.0	7	23	ABJ03861	Human VEGF-targete
12	36	100.0	7	23	ABJ03862	Human VEGF-targete
13	36	100.0	7	23	ABJ03863	Human VEGF-targete
14	36	100.0	7	23	ABJ03866	Human VEGF-targete
15	36	100.0	7	23	ABJ03868	Human VEGF-targete
16	36	100.0	7	23	ABJ03870	Human VEGF-targete
17	36	100.0	7	23	ABJ03871	Human VEGF-targete
18	36	100.0	7	23	ABJ03873	Human VEGF-targete
19	36	100.0	7	23	ABJ03897	Human VEGF-targete
20	36	100.0	7	23	ABJ03906	Human VEGF-targete
21	36	100.0	7	23	ABJ03912	Human VEGF-targete
22	36	100.0	7	23	ABJ03915	Human VEGF-targete
23	36	100.0	7	23	ABJ03917	Human VEGF-targete
24	36	100.0	7	23	ABJ03920	Human VEGF-targete
25	36	100.0	7	23	ABJ03939	Rat VEGF-targeted
26	36	100.0	7	23	ABB80809	Human ER-alpha loc
27	36	100.0	7	23	ABB80811	Human ER-alpha loc
28	36	100.0	7	23	ABB80809	Zinc finger protei
29	36	100.0	7	23	ABB80815	Zinc finger protei
30	36	100.0	7	23	ABB80836	Zinc finger protei
31	36	100.0	7	23	ABB80848	Zinc finger protei
32	36	100.0	7	23	ABB80863	Zinc finger protei
33	36	100.0	7	23	ABP48205	Zinc finger protei
34	36	100.0	7	23	ABP48208	Zinc finger protei
35	36	100.0	7	23	ABP48220	Zinc finger protei
36	36	100.0	7	23	ABP48232	Zinc finger protei
37	36	100.0	7	23	ABP48235	Zinc finger protei
38	36	100.0	7	23	ABP48250	Zinc finger protei
39	36	100.0	7	23	ABP48479	Zinc finger protei
40	36	100.0	7	23	ABP48484	Zinc finger protei
41	36	100.0	7	23	ABP48485	Zinc finger protei
42	36	100.0	7	23	ABP48487	Zinc finger protei
43	36	100.0	7	23	ABP48487	Zinc finger protei
44	36	100.0	7	23	ABP48515	Zinc finger protei
45	36	100.0	7	23	ABP48623	Zinc finger protei

ALIGNMENTS

RESULT 1
AAE08729
ID AAE08729 standard; peptide; 7 AA.
XX
AC AAE08729;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human KCA4 protein F3 recognition helix.
XX
KW Human; KCA4; EPO; molecular target; zinc finger protein; ZFP;
KW cellular process; signal transduction; drug-screening.
XX
OS Homo sapiens.
PN WO200159450-A2.
XX 16-AUG-2001.
XX
PF 08-FEB-2001; 2001WO-US04301.
XX
PR 08-FEB-2000; 2000US-0181117.
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX
XX Case C;
XX
XX WPI; 2001-522491/57.
XX
XX Screening compound for interaction with molecular target by contacting
XX compound with cells, comprising exogenous zinc finger protein that
XX modulates expression of target, and determining values of properties of
XX cells

XX Example 10; Page 73; 99pp; English.
 PS The invention relates to a method of screening a compound for interaction
 CC with a molecular target. The method involves contacting first and
 CC second cells with the compound and determining the values of properties
 CC of the compound. The second cell comprises an exogenous zinc finger
 CC protein (ZFP) that modulates the expression of the molecular target, or
 CC isolating membranes from cell comprising ZFP. The methods allow for high
 CC throughput screening of candidate compound and reduces the incidence of
 CC false positives. The methods are useful for screening a compound for
 CC interaction with a molecular target or for screening a compound for its
 CC effect on a cellular process. The method is useful for testing a compound
 CC for its capacity to transduce a signal to the molecular target or its
 CC capacity to block transduction of a signal through the molecular target,
 CC and for performing biochemical drug-screening assays. The present
 CC sequence is human Kc4 protein recognition helix used in the
 CC exemplification of the invention.
 XX
 SQ Sequence 7 AA;

Query Match 100.0%; Score 36; DB 22; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLRSR 7
 DB 1 RSDHLRSR 7

RESULT 2
 AAB84235
 ID AAB84235 standard; peptide; 7 AA.
 XX
 AC AAB84235;
 XX
 DT 06-AUG-2001 (first entry)
 XX
 DE Zinc protein recognition helix SBS3 for target DNA triplet GGG.
 XX
 KW Phenotype associated gene; zinc finger protein; cancer; nephritis;
 KW prostate hypertrophy; hematopoiesis; osteoporosis; obesity;
 KW cardiovascular disease; diabetes.

OS Synthetic.
 XX WO200140798-A2.
 PN
 PD 07-JUN-2001.
 XX
 PF 06-DEC-2000; 2000WO-US33086.
 XX
 PR 06-DEC-1999; 99US-0456100.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 PI Case CC, Liu Q, Rebar EJ;
 XX
 DR WPI; 2001-374953/39.
 XX

XX Identifying genes associated with selected phenotype for research
 PT purposes, involves culturing cells transduced with nucleic acid
 PT encoding zinc finger proteins and assaying cells exhibiting selected
 PT phenotype -

PS Example 1; Page 36; 58pp; English.

XX The specification describes a method for identifying genes associated
 CC with a selected phenotype. The method involves providing a library of
 CC nucleotide sequences encoding partially randomized zinc finger proteins,
 CC transducing cells with expression vectors, each comprising a sequence
 CC from the library, culturing the cells for expressing the zinc finger
 CC protein, assaying the cells for selected phenotype, and identifying the

CC gene of interest, in cells exhibiting the phenotype. The method is useful
 CC for identifying a gene or genes associated with a selected phenotype such
 CC as the one related to cancer, nephritis, prostate hypertrophy,
 CC hematopoiesis, osteoporosis, obesity, cardiovascular disease or diabetes.
 CC The method is useful in academic laboratories, in the biotechnological
 CC industries, and in pharmaceutical, genomic, agricultural and chemical
 CC companies. AAB84233-44 represent recognition helices of zinc finger
 CC proteins, which recognise different DNA triplets.
 XX

SQ Sequence 7 AA;

Query Match 100.0%; Score 36; DB 22; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLRSR 7
 DB 1 RSDHLRSR 7

RESULT 3
 AAB84247
 ID AAB84247 standard; peptide; 7 AA.
 XX

AC AAB84247;

DT 06-AUG-2001 (first entry)

DE Amino acid sequence of finger 3 of a zinc finger protein.

XX Phenotype associated gene; zinc finger protein; cancer; nephritis;
 KW prostate hypertrophy; hematopoiesis; osteoporosis; obesity;
 KW cardiovascular disease; diabetes.

OS Synthetic.

XX WO200140798-A2.

PN 07-JUN-2001.

XX 06-DEC-2000; 2000WO-US33086.

XX 06-DEC-1999; 99US-0456100.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Case CC, Liu Q, Rebar EJ;

XX WPI; 2001-374953/39.

XX Identifying genes associated with selected phenotype for research
 PT purposes, involves culturing cells transduced with nucleic acid
 PT encoding zinc finger proteins and assaying cells exhibiting selected
 PT phenotype -

PS Example 1; Page 41; 58pp; English.

XX The specification describes a method for identifying genes associated
 CC with a selected phenotype. The method involves providing a library of
 CC nucleotide sequences encoding partially randomized zinc finger proteins,
 CC transducing cells with expression vectors, each comprising a sequence
 CC from the library, culturing the cells for expressing the zinc finger
 CC protein, assaying the cells for selected phenotype, and identifying the
 CC gene of interest, in cells exhibiting the phenotype. The method is useful
 CC for identifying a gene or genes associated with a selected phenotype such
 CC as the one related to cancer, nephritis, prostate hypertrophy,
 CC hematopoiesis, osteoporosis, obesity, cardiovascular disease or diabetes.
 CC The method is useful in academic laboratories, in the biotechnological
 CC industries, and in pharmaceutical, genomic, agricultural and chemical
 CC companies. The present sequence represents a finger of a zinc finger
 CC protein.

XX Sequence 7 AA;

Query Match 100.0%; Score 36; DB 22; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
 |||||
 Db 1 RSDHLSR 7

RESULT 4
 ABB83583
 ID ABB83583 standard; peptide; 7 AA.
 XX
 AC ABB83583;
 XX
 DT 27-SEP-2002 (first entry)
 XX
 DE F3 zinc finger for target sequence ZFP 7.
 XX
 KW Zinc finger; stress tolerance; pathogen resistance;
 KW agrochemical.
 XX
 OS Unidentified.
 XX
 PN WO200257294-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 22-JAN-2002; 2002WO-US01906.
 XX
 PR 22-JAN-2001; 2001US-263445P.
 XX
 PR 11-MAY-2001; 2001US-290716P.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 PI Jamieson A, Li G;
 XX
 DR WPI; 2002-566792/60.
 XX
 PT Modified plant zinc finger protein for modulating gene expression in a
 PT plant cell comprises zinc fingers that bind to a target site -
 XX
 PS Example 4; Page 42; 50pp; English.
 XX
 CC The present invention relates to a modified plant zinc finger
 CC protein. This zinc finger protein is used to modulated gene
 CC expression in a plant cell. Nucleic acid encoding the zinc finger is
 CC expressed in plant cells to produce a plant with an altered phenotype
 CC relative to the wild-type plant. The altered phenotype is high in
 CC nutritional value, yield, stress tolerance, pathogen resistance,
 CC resistance to agrochemicals, production of pharmaceutical compounds or
 CC production of industrial chemicals. The present sequence is
 CC a zinc finger protein sequence that is attracted to a ZFP target
 CC sequence.
 XX
 SQ Sequence 7 AA;

Query Match 100.0%; Score 36; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
 |||||
 Db 1 RSDHLSR 7

RESULT 5
 ABB835821
 ID ABB835821 standard; Peptide; 7 AA.
 XX
 AC ABB835821;
 XX

DT 25-SEP-2002 (first entry)
 XX
 DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 64.
 XX
 KW Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
 KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
 KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnery;
 KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
 KW osteopathic; antiinfertility.
 XX
 OS Homo sapiens.
 XX
 PN WO200246412-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 06-DEC-2001; 2001WO-US46861.
 XX
 PR 07-DEC-2000; 2000US-0733604.
 PR 12-DEC-2000; 2000US-0736083.
 PR 30-APR-2001; 2001US-0846033.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
 PI Jarvis E;
 XX
 DR WPI; 2002-527918/56.
 XX
 PT New zinc finger protein that binds to target site in vascular
 PT endothelial growth factor gene, useful for modulating expression of the
 PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
 PT ulcer
 XX
 PS Claim 4; Page 102; 195pp; English.
 XX
 CC The present invention relates to a zinc finger protein that binds to a
 CC target site in one or more vascular endothelial growth factor (VEGF)
 CC genes. The protein is useful for modulating expression of a VEGF gene,
 CC thereby regulating angiogenesis and vasculogenesis. This can be used to
 CC treat atherosclerosis, ischemia, arthritis, wounds, ulcers, tumours,
 CC diabetic retinopathy or psoriasis. The present sequence is a peptide
 CC shown in the invention.
 XX
 SQ Sequence 7 AA;

Query Match 100.0%; Score 36; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
 |||||
 Db 1 RSDHLSR 7

RESULT 6
 ABB83825
 ID ABB83825 standard; Peptide; 7 AA.
 XX
 AC ABB83825;
 XX
 DT 25-SEP-2002 (first entry)
 XX
 DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 68.
 XX
 KW Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
 KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
 KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnery;
 KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
 KW osteopathic; antiinfertility.
 XX
 OS Homo sapiens.
 XX

PN WO200246412-A2.
XX 13-JUN-2002.
XX 06-DEC-2001; 2001WO-US46861.
XX 07-DEC-2000; 2000US-0733604.
PR 12-DEC-2000; 2000US-0736083.
PR 30-APR-2001; 2001US-0846033.
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX
PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
PI Jarvis E;
XX WPI; 2002-527918/56.
XX
XX New zinc finger protein that binds to target site in vascular
PT endothelial growth factor gene, useful for modulating expression of the
PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
PT ulcer
XX
PS Claim 4; Page 102; 195pp; English.
XX
XX The present invention relates to a zinc finger protein that binds to a
CC target site in one or more vascular endothelial growth factor (VEGF)
CC genes. The protein is useful for modulating expression of a VEGF gene,
CC thereby regulating angiogenesis and vasculogenesis. This can be used to
CC treat atherosclerosis, ischemia, arthritis, wounds, ulcers, tumours,
CC diabetic retinopathy or psoriasis. The present sequence is a peptide
CC shown in the invention.
XX
SQ Sequence 7 AA;
XX
XX Query Match 100.0%; Score 36; DB 23; Length 7;
XX Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RSDHLSR 7
DB 1 RSDHLSR 7
IIIIII
RESULT 7
ABJ03848
ID ABJ03848 standard; Peptide; 7 AA.
XX
AC ABJ03848;
XX
DT 25-SEP-2002 (first entry)
XX
DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 91.
XX
XX Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnery;
KW antulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
KW osteopathic; antiinfertility.
XX
OS Homo sapiens.
XX
XX WO200246412-A2.
XX 13-JUN-2002.
XX
XX 06-DEC-2001; 2001WO-US46861.
XX
XX 07-DEC-2000; 2000US-0733604.
PR 12-DEC-2000; 2000US-0736083.
PR 30-APR-2001; 2001US-0846033.
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX
PA Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
PI Jarvis E;
XX WPI; 2002-527918/56.
XX
XX New zinc finger protein that binds to target site in vascular
PT endothelial growth factor gene, useful for modulating expression of the
PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
PT ulcer
XX
PS Claim 4; Page 102; 195pp; English.
XX
XX The present invention relates to a zinc finger protein that binds to a
CC target site in one or more vascular endothelial growth factor (VEGF)
CC genes. The protein is useful for modulating expression of a VEGF gene,
CC thereby regulating angiogenesis and vasculogenesis. This can be used to
CC treat atherosclerosis, ischemia, arthritis, wounds, ulcers, tumours,
CC diabetic retinopathy or psoriasis. The present sequence is a peptide
CC shown in the invention.
XX
SQ Sequence 7 AA;
XX
XX Query Match 100.0%; Score 36; DB 23; Length 7;
XX Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RSDHLSR 7
DB 1 RSDHLSR 7
IIIIII
RESULT 8
ABJ03858
ID ABJ03858 standard; Peptide; 7 AA.
XX
AC ABJ03858;
XX
DT 25-SEP-2002 (first entry)
XX
DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 101.
XX
XX Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnery;
KW antulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
KW osteopathic; antiinfertility.
XX
OS Homo sapiens.
XX
XX WO200246412-A2.
XX 13-JUN-2002.
XX
XX 06-DEC-2001; 2001WO-US46861.
XX
XX 07-DEC-2000; 2000US-0733604.
PR 12-DEC-2000; 2000US-0736083.
PR 30-APR-2001; 2001US-0846033.
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX
PA Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
PI Jarvis E;
XX WPI; 2002-527918/56.
XX
XX New zinc finger protein that binds to target site in vascular
PT endothelial growth factor gene, useful for modulating expression of the
PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
PT ulcer
XX
PS Claim 4; Page 102; 195pp; English.
XX

PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
PI Jarvis E;
XX WPI; 2002-527918/56.
XX
XX New zinc finger protein that binds to target site in vascular
PT endothelial growth factor gene, useful for modulating expression of the
PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
PT ulcer
XX
PS Claim 4; Page 102; 195pp; English.
XX
XX The present invention relates to a zinc finger protein that binds to a
CC target site in one or more vascular endothelial growth factor (VEGF)
CC genes. The protein is useful for modulating expression of a VEGF gene,
CC thereby regulating angiogenesis and vasculogenesis. This can be used to
CC treat atherosclerosis, ischemia, arthritis, wounds, ulcers, tumours,
CC diabetic retinopathy or psoriasis. The present sequence is a peptide
CC shown in the invention.
XX
SQ Sequence 7 AA;
XX
XX Query Match 100.0%; Score 36; DB 23; Length 7;
XX Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RSDHLSR 7
DB 1 RSDHLSR 7
IIIIII
RESULT 8
ABJ03858
ID ABJ03858 standard; Peptide; 7 AA.
XX
AC ABJ03858;
XX
DT 25-SEP-2002 (first entry)
XX
DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 101.
XX
XX Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnery;
KW antulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
KW osteopathic; antiinfertility.
XX
OS Homo sapiens.
XX
XX WO200246412-A2.
XX 13-JUN-2002.
XX
XX 06-DEC-2001; 2001WO-US46861.
XX
XX 07-DEC-2000; 2000US-0733604.
PR 12-DEC-2000; 2000US-0736083.
PR 30-APR-2001; 2001US-0846033.
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX
PA Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
PI Jarvis E;
XX WPI; 2002-527918/56.
XX
XX New zinc finger protein that binds to target site in vascular
PT endothelial growth factor gene, useful for modulating expression of the
PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
PT ulcer
XX
PS Claim 4; Page 102; 195pp; English.
XX


```

XX DT 25-SEP-2002 (first entry)
XX DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 104.
XX DE
XX DE
XX KW Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
XX KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
XX KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnerary;
XX KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
XX KW osteopathic; antiinfertility.
XX OS Homo sapiens.
XX XX
XX PN WO200246412-A2.
XX PD 13-JUN-2002.
XX PF 06-DEC-2001; 2001WO-US46861.
XX PR 07-DEC-2000; 2000US-0733604.
XX PR 12-DEC-2000; 2000US-0736083.
XX PR 30-APR-2001; 2001US-0846033.
XX PA (SANG-) SANGAMO BIOSCIENCES INC.
XX PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
XX PI Jarvis E;
XX XX
XX DR WPI; 2002-527918/56.
XX XX
XX PT New zinc finger protein that binds to target site in vascular
XX PT endothelial growth factor gene, useful for modulating expression of the
XX PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
XX PT ulcer
XX PS Claim 4; Page 102; 195pp; English.
XX XX
XX CC The present invention relates to a zinc finger protein that binds to a
XX CC target site in one or more vascular endothelial growth factor (VEGF)
XX CC genes. The protein is useful for modulating expression of a VEGF gene,
XX CC thereby regulating angiogenesis and vasculogenesis. This can be used to
XX CC treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
XX CC diabetic retinopathy or psoriasis. The present sequence is a peptide
XX CC shown in the invention.
XX SQ Sequence 7 AA;
XX
XX Query Match 100.0%; Score 36; DB 23; Length 7;
XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 RSDHLSR 7
XX Db | | | | | |
XX 1 RSDHLSR 7
XX
XX RESULT 13
XX ABJ03863
XX ID ABJ03863 standard; Peptide; 7 AA.
XX XX
XX AC ABJ03863;
XX XX
XX DT 25-SEP-2002 (first entry)
XX XX
XX DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 106.
XX XX
XX KW Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
XX KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
XX KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnerary;
XX KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
XX KW osteopathic; antiinfertility.
XX OS Homo sapiens.
XX XX
XX PN WO200246412-A2.
XX PD 13-JUN-2002.
XX PF 06-DEC-2001; 2001WO-US46861.
XX PR 07-DEC-2000; 2000US-0733604.
XX PR 12-DEC-2000; 2000US-0736083.
XX PR 30-APR-2001; 2001US-0846033.
XX PA (SANG-) SANGAMO BIOSCIENCES INC.
XX PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
XX PI Jarvis E;
XX XX
XX DR WPI; 2002-527918/56.
XX XX
XX PT New zinc finger protein that binds to target site in vascular
XX PT endothelial growth factor gene, useful for modulating expression of the
XX PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
XX PT ulcer
XX PS Claim 4; Page 102; 195pp; English.
XX XX
XX CC The present invention relates to a zinc finger protein that binds to a
XX CC target site in one or more vascular endothelial growth factor (VEGF)
XX CC genes. The protein is useful for modulating expression of a VEGF gene,
XX CC thereby regulating angiogenesis and vasculogenesis. This can be used to
XX CC treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
XX CC diabetic retinopathy or psoriasis. The present sequence is a peptide
XX CC shown in the invention.
XX SQ Sequence 7 AA;
XX
XX Query Match 100.0%; Score 36; DB 23; Length 7;
XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 RSDHLSR 7
XX Db | | | | | |
XX 1 RSDHLSR 7
XX
XX RESULT 12
XX ABJ03862
XX ID ABJ03862 standard; Peptide; 7 AA.
XX XX
XX AC ABJ03862;
XX XX
XX DT 25-SEP-2002 (first entry)
XX XX
XX DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 105.
XX XX
XX KW Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
XX KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
XX KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnerary;
XX KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
XX KW osteopathic; antiinfertility.
XX OS Homo sapiens.
XX

```


XX Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
PI Jarvis E;
XX WPI; 2002-527918/56.
XX New zinc finger protein that binds to target site in vascular
PT endothelial growth factor gene, useful for modulating expression of the
PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
PT ulcer -
XX Claim 4; Page 102; 195pp; English.
XX The present invention relates to a zinc finger protein that binds to a
CC target site in one or more vascular endothelial growth factor (VEGF)
CC genes. The protein is useful for modulating expression of a VEGF
CC gene, thereby regulating angiogenesis and vasculogenesis. This can be used to
CC treat atherosclerosis, ischemia, arthritis, wounds, ulcers, tumours,
CC diabetic retinopathy or psoriasis. The present sequence is a peptide
CC shown in the invention.
XX Sequence 7 AA;
PS Query Match 100.0%; Score 36; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RSDHLSR 7
Db 1 RSDHLSR 7
RESULT 14
ABJ03866
ID ABJ03866 standard; Peptide; 7 AA.
XX AC ABJ03866;
XX 25-SEP-2002 (first entry)
DT Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 109.
DE Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
XX diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnerary;
KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
KW osteopathic; antiinfertility.
XX Homo sapiens.
OS
XX WO200246412-A2.
PN 13-JUN-2002.
XX 06-DEC-2001; 2001WO-US46861.
XX 07-DEC-2000; 2000US-0733604.
PR 12-DEC-2000; 2000US-0736083.
PR 30-APR-2001; 2001US-0846033.
XX (SANG-) SANGAMO BIOSCIENCES INC.
PA Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
PI Jarvis E;
XX WPI; 2002-527918/56.
XX New zinc finger protein that binds to target site in vascular
PT endothelial growth factor gene, useful for modulating expression of the
PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
PT ulcer -
XX Claim 4; Page 102; 195pp; English.

XX The present invention relates to a zinc finger protein that binds to a
CC target site in one or more vascular endothelial growth factor (VEGF)
CC genes. The protein is useful for modulating expression of a VEGF gene,
CC thereby regulating angiogenesis and vasculogenesis. This can be used to
CC treat atherosclerosis, ischemia, arthritis, wounds, ulcers, tumours,
CC diabetic retinopathy or psoriasis. The present sequence is a peptide
CC shown in the invention.
XX Sequence 7 AA;
PS Query Match 100.0%; Score 36; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RSDHLSR 7
Db 1 RSDHLSR 7
RESULT 15
ABJ03868
ID ABJ03868 standard; Peptide; 7 AA.
XX AC ABJ03868;
XX 25-SEP-2002 (first entry)
DT Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 111.
DE Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
XX diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnerary;
KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
KW osteopathic; antiinfertility.
XX Homo sapiens.
OS
XX WO200246412-A2.
PN 13-JUN-2002.
XX 06-DEC-2001; 2001WO-US46861.
XX 07-DEC-2000; 2000US-0733604.
PR 12-DEC-2000; 2000US-0736083.
PR 30-APR-2001; 2001US-0846033.
XX (SANG-) SANGAMO BIOSCIENCES INC.
PA Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
PI Jarvis E;
XX WPI; 2002-527918/56.
XX New zinc finger protein that binds to target site in vascular
PT endothelial growth factor gene, useful for modulating expression of the
PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
PT ulcer -
XX Claim 4; Page 102; 195pp; English.
XX The present invention relates to a zinc finger protein that binds to a
CC target site in one or more vascular endothelial growth factor (VEGF)
CC genes. The protein is useful for modulating expression of a VEGF gene,
CC thereby regulating angiogenesis and vasculogenesis. This can be used to
CC treat atherosclerosis, ischemia, arthritis, wounds, ulcers, tumours,
CC diabetic retinopathy or psoriasis. The present sequence is a peptide
CC shown in the invention.
XX Sequence 7 AA;
PS Query Match 100.0%; Score 36; DB 23; Length 7;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
|||||||
Db 1 RSDHLSR 7

Search completed: July 15, 2003, 12:53:01
Job time : 34.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:49:31 ; Search time 6.66667 Seconds
(without alignments)
43.550 Million cell updates/sec

Title: US-10-006-069A-247

Perfect score: 36

Sequence: 1 TSGHLSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	36	100.0	220	1 NRG2_YEAST	P38082 saccharomyc
2	33	91.7	231	1 NRG1_YEAST	Q03125 saccharomyc
3	32	88.9	214	1 VTI1_SCHPO	P78768 schizosacch
4	31	86.1	352	1 TPO_CANFA	P42705 canis fami
5	31	86.1	423	1 V107_VAGCC	P20501 vaccinia vi
6	31	86.1	423	1 V107_VAGCV	P12926 vaccinia vi
7	31	86.1	530	1 ESR2_MOUSE	O08537 mus musculu
8	31	86.1	530	1 ESR2_RAT	Q62986 rattus norv
9	31	86.1	747	1 ORPB_HUMAN	Q9bxb4 homo sapien
10	30	83.3	783	1 RIN1_HUMAN	Q13671 homo sapien
11	29	80.6	105	1 RL36_TRITHM	Q9hfr7 trichoderma
12	29	80.6	119	1 HTS3_RHOSH	Q53158 rhodobacter
13	29	80.6	271	1 HEXP_LEIMA	Q04832 leishmania
14	28	77.8	170	1 CNBP_MOUSE	P53996 mus musculu
15	28	77.8	172	1 CNBP_CHICK	O42395 gallus gall
16	28	77.8	177	1 CNBP_HUMAN	P20694 homo sapien
17	28	77.8	251	1 GLO2_BUCAI	P57336 buchiera ap
18	28	77.8	268	1 CAPB_SCHPO	Q9hqp5 schizosacch
19	28	77.8	321	1 ANX5_CHICK	P17153 gallus gall
20	28	77.8	324	1 ANX9_DROME	P22464 drosophila
21	28	77.8	391	1 GPD1_YEAST	Q00055 saccharomyc
22	28	77.8	463	1 ANX7_MOUSE	Q07076 mus musculu
23	28	77.8	466	1 ANX7_HUMAN	P20073 homo sapien
24	28	77.8	466	1 ARLY_BRUME	Q8vj17 brucella me
25	28	77.8	475	1 SVC_AERPE	Q3ydk6 aeropyrum p
26	28	77.8	503	1 ANXB_BOVIN	P27214 bos taurus
27	28	77.8	503	1 ANXB_MOUSE	P37384 mus musculu
28	28	77.8	503	1 ANXB_RABIT	P33477 oryctolagus
29	28	77.8	505	1 ANXB_HUMAN	P50995 homo sapien
30	28	77.8	512	1 ANX7_XENLA	Q92125 xenopus lae
31	28	77.8	618	1 ANX6_BOVIN	P79134 bos taurus
32	28	77.8	625	1 DXS_HAETIN	P45205 haemophilus
33	28	77.8	644	1 ARNT_DROME	O15945 drosophila

ALIGNMENTS

RESULT 1

NRG2_YEAST
ID NRG2_YEAST STANDARD; PRT; 220 AA.
AC P38082;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable transcriptional regulator NRG2.
GN NRG2 OR YBR066C OR YBR0616.
OS Saccharomyces cerevisiae (Eaker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE OF 1-187 FROM N.A.
RC STRAIN=S288C;
RA Feldmann H., Mannhaupt G., Schwarzlose C., Vetter I.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 124-220 FROM N.A.
RC STRAIN=S288C;
RA Domdey H., Gassenhuber H., Obermaier B., Piravandi E.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSCRIPTIONAL REPRESSOR (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: STRONG, TO NRG1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
DR EMBL; Z35935; CAA85010.1;
DR FIR; S45927; S45927.
DR TRANSFAC; T03517;
DR SGD; S0000270; NRG2.
DR InterPro; IPR000822; znf_C2H2.
DR Pfam; PF00096; znf_C2H2_2.
DR PRINTS; PR00048; ZINC_FINGER.
DR SMART; SM00355; znf_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
KW DNA-binding; Nuclear protein; Zinc-finger; Metal-binding; Repressor;
KW Transcription regulation; Repeat.
FT ZN_FING 153 175 C2H2-TYPE.
FT ZN_FING 181 205 C2H2-TYPE.
SQ SEQUENCE 220 AA; 25009 MW; 8EA18F326910F1FB CRC64;

Query Match 100.0% Score 36; DB 1; Length 220;
Best Local Similarity 100.0%; Pred. No. 0.95; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TSGHLSR 7

Db 164 TSGHLSR 170
|||||

RESULT 2

NRGL_YEAST STANDARD; PRT; 231 AA.
AC Q03125;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Transcriptional regulator NRGI (zinc finger protein MSS1).
GN NRGI OR MSS1 OR YDR043C OR YD5112.OIC.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RA Oliver K., Harris D., Barrell B., Rajandream M.A., Walsh S.V.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AN20-5B;
RA Park S.H., Hwang H.J., Kang H.S.;
RT "MSS1 gene, encoding a zinc-finger motif, is involved in glucose
RT repression and STAL0 repression of glucoamylase gene in Saccharomycetes
RT cerevisiae.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR INVOLVED IN REGULATION OF
CC GLUCOSE REPRESSION. BINDS TO UAS-1 IN THE STAL PROMOTER.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: STRONG, TO NRG2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z49812; CAA09965.1; -
CC EMBL; AF013167; AAD01510.1; -
CC SGD; S0002450; NRGI.
CC InterPro: IPR000822; znf_C2H2.
CC Pfam: PF00096; zf_C2H2; 2.
CC PRINTS; PR00048; ZINC_FINGER.
CC ProDom; PD000003; znf_C2H2; 1.
CC SMART; SM00355; znf_C2H2; 2.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
CC DNA-binding; Nuclear protein; zinc-finger; Metal-binding; Repressor;
KW Transcription regulation; Repeat.
FT ZN_FING 174 196 C2H2-TYPE.
FT ZN_FING 202 226 C2H2-TYPE.
SQ SEQUENCE 231 AA; 26743 MW; 50F7E7ED4690D109 CRC64;

Query Match 91.7%; Score 33; DB 1; Length 231;
Best Local Similarity 85.7%; Pred. No. 4.5;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
|||||
Db 185 TSGHLAR 191

RESULT 3

VTIL_SCHPO STANDARD; PRT; 214 AA.
AC P78768; O43039;
DT 15-JUL-1998 (Rel. 36, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vesicle transport v-SNARE protein VTIL homolog.
GN SPBC3B9.10
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Bigham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens J., Vansteleels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Pohl T.M.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong I., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=98162722; PubMed=9501991;
RA Yoshioke S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT cDNAs.";
RL DNA Res. 4:363-369(1997).
CC -1- FUNCTION: V-SNARE THAT MEDIATES VESICLE TRANSPORT PATHWAYS THROUGH
CC INTERACTIONS WITH T-SNARES ON THE TARGET MEMBRANE. THESE
CC INTERACTIONS ARE PROPOSED TO MEDIATE ASPECTS OF THE SPECIFICITY OF
CC VESICLE TRAFFICKING AND TO PROMOTE FUSION OF THE LIPID BILAYERS
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE VTIL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL022070; CAA17790.1; -
CC EMBL; D89116; BAAL3778.1; -
CC InterPro: IPR000727; T-SNARE.
CC SMART; SM00397; t-SNARE; 1.
KW Transmembrane; Transport; Protein transport.
FT TRANSMEM 192 212 POTENTIAL.
FT CONFLICT 135 135 O -> H (IN REF. 2).
SQ SEQUENCE 214 AA; 24743 MW; 641A57E4D7E52A6F CRC64;

Query Match 88.9%; Score 32; DB 1; Length 214;
Best Local Similarity 85.7%; Pred. No. 6.8;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
| | | | |
Db 169 TSGHLDR 175

RESULT 4

TPO_CANFA STANDARD; PRT; 352 AA.
AC P42705;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
DE (C-MPL ligand) (ML) (Megakaryocyte growth and development factor)
DE (MGDF).
GN THPO OR TPO.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-44.
RC TISSUE=Kidney;
RX MEDLINE=94291201; PubMed=8020099;
RA Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,
RA Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
RA Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Merewether L.A.,
RA Clogson C., Hsu E., Hsok M.M., Hornkohl A., Choi E., Pangelinan M.,
RA Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C.,
RA Shutter J., Chute H., Basu R., Selander L., Trollinger D., Sieu L.,
RA Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
RA Garcia A., Xu W., Del Castillo J., Biron J., Cole S., Hu M.C.-T.,
RA Pacifici R., Ponting I., Sarlis C., Wen D., Yung Y.P., Lin H.,
RA Bosselman R.A.;
RT "Identification and cloning of a megakaryocyte growth and development
RT factor that is a ligand for the cytokine receptor Mpl.";
RL Cell 77:1117-1117(1994).
CC -1- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
CC AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
CC CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
CC MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
CC TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
DR InterPro; IPR001323; EPO_TPO.
DR Pfam; PF00758; thrombopoietin.
DR PRINTS; PR01485; THROMBOPTN.
DR PROSITE; PS00817; EPO_TPO; 1.
KW Cytokine; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 23
FT CHAIN 24 352 THROMBOPOIETIN.
FT DISULFID 28 172 POTENTIAL.
FT DISULFID 50 106 POTENTIAL.
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 352 AA; 37641 MW; 024F3B41B061FDB8 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 352;

Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
| | | | |
Db 243 TPGHLSR 249

RESULT 5

VI07_VACCC STANDARD; PRT; 423 AA.
AC P20501;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein I7.
GN I7L.
OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10249;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021027; PubMed=2219722;
RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RT "The complete DNA sequence of vaccinia virus.";
RL Virology 179:247-266(1990).
RN [2]
RP COMPLETE GENOME.
RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RT "Appendix to 'The complete DNA sequence of vaccinia virus'.";
RL Virology 179:517-563(1990).
CC -1- SIMILARITY: BELONGS TO THE POXVIRUSES I7 FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M35027; AAA48063.1; -
CC PIR; C42511; C42511.
CC NEROPS; C57.001; -
CC InterPro; IPR004970; Peptidase_C57.
CC Pfam; PF03290; Peptidase_C57; 1.
KW Late protein.
SQ SEQUENCE 423 AA; 49039 MW; 633F501637FD80D3 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 423;

Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGHLSR 7
| | | | |
Db 74 SGHLSR 79

RESULT 6

VI07_VACCV STANDARD; PRT; 423 AA.
AC P12926;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein I7.
GN I7L.
OS Vaccinia virus (strain WR).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10254;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88215015; PubMed=2835495;
RA Schmitt J.F.C., Stunnenberg H.G.;
RT "Sequence and transcriptional analysis of the vaccinia virus HindIII

RT I fragment.";
 RL J. Virol. 62:1889-1897(1988).
 RN [2]
 RP SEQUENCE OF 1-215 FROM N.A.
 RX MEDLINE-91134989; PubMed-1994576;
 RA Fathi Z., Condit R.C.;
 RT "Genetic and molecular biological characterization of a vaccinia
 RT virus temperature-sensitive complementation group affecting a virion
 RT component.";
 RL Virology 181:258-272(1991).
 CC -1- SIMILARITY: BELONGS TO THE POXVIRUSES 17 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J03399; AAB59809.1; -
 DR PIR; G29889; W2V217.
 DR PIR; A38497; A38497.
 DR InterPro; IPR004970; Peptidase_C57.
 DR Pfam; PF03290; Peptidase_C57; 1.
 DR Late protein.
 KW
 SQ SEQUENCE 423 AA; 49010 MW; 6E5BF25D7FD9FEE CRC64;
 Query Match 86.1%; Score 31; DB 1; Length 423;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SGHLSR 7
 Db 74 SGHLSR 79
 RESULT 7
 ESR2_MOUSE STANDARD; PRT; 530 AA.
 ID ESR2_MOUSE 008537; 070519;
 AC 008537; 035635; 070519;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Estrogen receptor beta (ER-beta).
 GN ESR2 OR NR3A2 OR ESRB.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE OF 46-530 FROM N.A., AND MUTAGENESIS.
 RC STRAIN-129/SV; TISSUE-Ovary;
 RX MEDLINE-97211383; PubMed-9058381;
 RA Tremblay G.B., Tremblay A., Copeland N.G., Gilbert D.J., Jenkins N.A.,
 RA Labrie F., Giguere V.;
 RT "Cloning, chromosomal localization, and functional analysis of the
 RT murine estrogen receptor beta.";
 RT Mol. Endocrinol. 11:353-365(1997).
 RN [2]
 RP SEQUENCE OF 46-530 FROM N.A., AND CHARACTERIZATION.
 RC TISSUE-Ovary;
 RX MEDLINE-97424142; PubMed-9280064;
 RA Pettersson K., Grandien K., Kuiper G.G.J.M., Gustafsson J.-A.;
 RT "Mouse estrogen receptor beta forms estrogen response element-binding
 RT heterodimers with estrogen receptor alpha.";
 RL Mol. Endocrinol. 11:1486-1496(1997).
 RN [3]
 RP SEQUENCE OF 1-60 FROM N.A.
 RC TISSUE-Ovary;
 RA Leygue E., Lu B., Dotzlaw H., Glor C., Watson P.H., Murphy L.C.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]

RP SEQUENCE OF 1-60 FROM N.A.
 RC STRAIN-C57BL/6J X 129; TISSUE-Ovary;
 RA Rosenfeld C.S., Lubahn D.B.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP CARBOHYDRATE-LINKAGE SITE SER-61, AND PHOSPHORYLATION SITE SER-61.
 RX MEDLINE-20450734; PubMed-1095228;
 RA Cheng X., Cole R.N., Zaia J., Hart G.W.;
 RT "Alternative O-glycosylation/O-phosphorylation of the murine estrogen
 RT receptor beta.";
 RL Biochemistry 39:11609-11620(2000).
 RN [6]
 RP ALTERNATIVE SPLICING.
 RC TISSUE-Ovary;
 RX MEDLINE-98348389; PubMed-9685228;
 RA Lu B., Leygue E., Dotzlaw H., Murphy L.J., Murphy L.C., Watson P.H.;
 RT "Estrogen receptor-beta mRNA variants in human and murine tissues.";
 RL Mol. Cell. Endocrinol. 138:199-203(1998).
 RN [7]
 RP TISSUE SPECIFICITY.
 RX MEDLINE-98268825; PubMed-9607809;
 RA Rosenfeld C.S., Ganjam V.K., Taylor J.A., Yuan X., Stiehr J.R.,
 RA Hardy M.P., Lubahn D.B.;
 RT "Transcription and translation of estrogen receptor-beta in the male
 RT reproductive tract of estrogen receptor-alpha knock-out and wild-type
 RT mice.";
 RL Endocrinology 139:2982-2987(1998).
 CC -1- FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
 CC ER-ALPHA, AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
 CC ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER.
 CC -1- SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
 CC ALPHA.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 5 ISOFORMS: BETA-1 (SHOWN HERE),
 CC BETA-2, BETA-5, BETA-5A AND BETA-6; ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN PROSTATE, OVARY, LEYDIG CELLS AND
 CC IN EPITHELIUM OF THE EFFERENT DUCTULES AND OF THE INITIAL SEGMENT
 CC OF THE EPIDIDYMIS.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NR3 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U81451; AAB51132.1; -
 DR EMBL; AT000220; CAA03949.1; -
 DR EMBL; AF067422; AAC17919.1; ALT_INIT.
 DR EMBL; AF063853; AAC16656.1; -
 DR HSSP; P03372; IHCO.
 DR TRANSFAC; T04649; -
 DR GlycoSuiteDB; O08537; -
 DR MGD; MGI:109392; Esr2.
 DR InterPro; IPR000536; Hormone_rec_lig.
 DR InterPro; IPR001723; Stchrnm_receptor.
 DR InterPro; IPR001628; Znf_C4steroid.
 DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STRDHORMONER.
 DR PRINTS; PR00047; STROIDFINGER.
 DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00430; HOLI; 1.
 DR SMART; SM00399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Steroid-binding; Alternative splicing; Glycoprotein;

KW Phosphorylation. 1 148
FT DOMAIN 149 214
FT ZN_FING 149 169
FT ZN_FING 185 209
FT DOMAIN 215 530
FT MOD_RES 61 61
FT MOD_RES 87 87
FT MOD_RES 105 105
FT MOD_RES 488 488
FT CARBOHYD 61 61
FT VARSPLIC 364 364
FT VARSPLIC 319 364
FT VARSPLIC 365 409
FT VARSPLIC 319 409
FT MUTAGEN 105 105
FT MUTAGEN 139 139
FT CONFLICT 47 47
FT CONFLICT 142 142
FT CONFLICT 200 200
FT CONFLICT 378 378
FT CONFLICT 412 412
FT CONFLICT 445 445
FT CONFLICT 511 511
SQ SEQUENCE 530 AA; 59012 MW; 90E73A161550ABE4 CRC64;
Query Match 86.1%; Score 31; DB 1; Length 530;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TSGHLS 6
Db 82 TSGHLS 87
RESULT 8
ESR2_RAT
ID ESR2_RAT STANDARD; PRT: 530 AA.
AC Q62986; 070195; 035785; 055015; 035784; Q9R185;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Estrogen receptor beta (ER-beta).
GN ESR2 OR NR3A2 OR ERBETA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RP SEQUENCE OF 46-530 FROM N.A. (ISOFORM BETA-1).
RC STRAIN-Sprague-Dawley; TISSUE=Prostate;
RX MEDLINE=96234066; PubMed=8650195;
RA Kuiper G.G.J.M., Enmark E., Peltö-Huikko M., Nilsson S., Gustafsson J.-A.;
RT "Cloning of a novel receptor expressed in rat prostate and ovary.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:5925-5930(1996).
RN [2]
RP SEQUENCE OF 46-530 FROM N.A. (ISOFORM BETA-2).
RC STRAIN=Wiistar; TISSUE=Ovary;
RX MEDLINE=98262932; PubMed=9600083;
RA Maruyama K., Endoh H., Sasaki-Iwaoka H., Kanou H., Shimaya E., Hashimoto S., Kato S., Kawashima H.;
RT "A novel isoform of rat estrogen receptor beta with 18 amino acid insertion in the ligand binding domain as a putative dominant negative regular of estrogen action.";
RL Biochem. Biophys. Res. Commun. 246:142-147(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM BETA-2).
RC STRAIN=Wiistar; TISSUE=Prostate;
Aldridge T.C.;
"tissue specific responses to estrogen: an explanation based on differential activation of multiple estrogen receptors with different estrogen response elements.";
Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE OF 46-530 FROM N.A. (ISOFORMS BETA-1; BETA-2; DELTA-3).
RC STRAIN-Sprague-Dawley;
RX MEDLINE=98151005; PubMed=9492041;
RA Petersen D.N., Tkalec G.T., Koza-Taylor P.H., Turi T.G., Brown T.A.;
RT "Identification of estrogen receptor beta2, a functional variant of estrogen receptor beta expressed in normal rat tissues.";
Endocrinology 139:1082-1092(1998).
[5]
RN PARTIAL SEQUENCE FROM N.A. (ISOFORM BETA-1/DELTA-4).
RC STRAIN-Sprague-Dawley; TISSUE=Brain;
RX Price R., Handa R.J.;
RT "A novel splice variant of estrogen receptor beta found in rat brain.";
Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF ER-ALPHA, AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER. ISOFORMS BETA-1/DELTA-3 AND BETA-2/DELTA-3 ARE UNABLE TO BIND DNA AND ACTIVATE TRANSCRIPTION DUE TO THE TRUNCATION OF THE DNA BINDING DOMAIN. ISOFORM BETA-2 SHOWS LOSS OF LIGAND BINDING AFFINITY AND SUPPRESSES ER-ALPHA AND ER-BETA-1 MEDIATED TRANSCRIPTIONAL ACTIVATION AND MAY ACT AS A DOMINANT NEGATIVE REGULATOR OF ESTROGEN ACTION.
-1- SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-ALPHA. CAN ALSO FORM HETERODIMERS BETWEEN ISOFORMS BETA-1 AND BETA-2.
-1- SUBCELLULAR LOCATION: Nuclear.
-1- ALTERNATIVE PRODUCTS: AT LEAST 5 ISOFORMS; BETA-1 (SHOWN HERE), BETA-1/DELTA-3, BETA-1/DELTA-4, BETA-2 AND BETA-2/DELTA-3; ARE PRODUCED BY ALTERNATIVE SPLICING.
-1- TISSUE SPECIFICITY: EXPRESSED IN PROSTATE, OVARY, LIVER, KIDNEY, FAT, BONE, BRAIN, UTERUS AND TESTIS.
-1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN, A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
-1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
NR3 SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: U57439; AAC52602.1;
EMBL: AB012721; BAA25431.1;
EMBL: AJ002602; CAA05631.1; ALT_INIT.
EMBL: AJ002603; CAA05632.1;
EMBL: AF042058; AAB97424.1;
EMBL: AF042059; AAB97425.1;
EMBL: AF042060; AAB97426.1;
EMBL: AF042061; AAB97427.1;
EMBL: AF161187; AAD47637.1;
HSSP: P03372; IHQ.
InterPro: IPR000536; Hormone_rec_lig.
InterPro: IPR001723; Steroid_receptor.
InterPro: IPR001628; Znf_C4steroid.
Pfam: PF00104; hormone_rec; 1.
Pfam: PF00105; zf-C4; 1.
PRINTS: PR00047; STERDORMONER.
PRINTS: PR00035; Znf_C4steroid; 1.
SMART: SM00430; HOLI; 1.
SMART: SM00399; Znf_C4; 1.
PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.

KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW zinc-finger; Steroid-binding; Alternative splicing; Phosphorylation.
FT DOMAIN 1 148 MODULATING.
FT DNA_BIND 149 214 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 149 169 C4-TYPE.
FT ZN_FING 185 209 C4-TYPE.
FT DOMAIN 215 530 STEROID-BINDING.
FT MOD_RES 87 87 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 488 488 PHOSPHORYLATION (BY SIMILARITY).
FT VARSPLIC 179 217 MISSING (IN ISOFORM BETA-1/DELTA-3 AND
FT VARSPLIC 219 318 ISOFORM BETA-2/DELTA-3).
FT VARSPLIC 364 364 MISSING (IN ISOFORM BETA-1/DELTA-4).
R -> RSSEDPHWHVAQMSAAPR (IN ISOFORM
BETA-2 AND ISOFORM BETA-2/DELTA-3).
L -> Q (IN REF. 3 AND 4).
L -> Q (IN REF. 3 AND 4).
P -> A (IN REF. 3 AND 4).
S -> P (IN REF. 3; CAA05631).
S -> P (IN REF. 3; CAA05631).
SQ SEQUENCE 530 AA; 59152 MW; 36F269D9FD773DA9 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 530;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
DB 82 TSGHLS 87

RESULT 9
ORPB_HUMAN STANDARD; PRT; 747 AA.
AC Q9BXB4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oxyesterol binding protein-related protein 11 (OSBP-related protein 11)
DE (ORP-11).
GN OSBP11 OR ORP11 OR OSBP12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21376257; PubMed=11483621;
RA Jaworski C.J., Moreira E., Li A., Lee R., Rodriguez I.R.;
RT "A family of 12 human genes containing oxysterol-binding domains.";
RL Genomics 78:185-196(2001).
RN [2]
RP SEQUENCE OF 1-315 FROM N.A.
RX MEDLINE=21376257; PubMed=11483621;
RA Lehto M., Laitinen S., Chinetti G., Johansson M., Ehnholm C.,
RA Staelen B., Ikonen E., Oikarinen V.M.;
RT "The OSBP-related protein family in humans.";
RL J. Lipid Res. 42:1203-1213(2001).
CC -1- TISSUE SPECIFICITY: Widely expressed.
CC -1- SIMILARITY: BELONGS TO THE OSBP FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; AF392454; AAL40667.1; -
CC DR EMBL; AF346292; AAK31141.1; -
CC DR Genew; HGNC:16397; OSBP11.
CC MIM: 606739; -
CC InterPro; IPR000648; Oxysterol_BP.

DR InterPro; IPR001849; PH.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF01237; Oxysterol_BP; 1.
DR SMART; SM00233; PH; 1.
DR PROSITE; PS01013; OSBP; FALSE_NEG.
DR PROSITE; PS00003; PH_DOMAIN; 1.
KW Lipid transport; Transport.
FT DOMAIN 58 155 PH.
SQ SEQUENCE 747 AA; 83643 MW; B443D3BDE8AE5FB6 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 747;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
DB 243 TSGHLS 248

RESULT 10
RINL_HUMAN STANDARD; PRT; 783 AA.
AC Q13671; Q00427; O15010;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras interaction/interference protein 1 (Ras inhibitor JC99).
GN RINL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP PRELIMINARY SEQUENCE FROM N.A.
RC TISSUE-Gliol cell.
RX MEDLINE=91187901; PubMed=1849280;
RA Colicelli J., Nicolette C., Birchmeier C., Rodgers L., Riggs M.,
RA Wigler M.;
RT "Expression of three mammalian cDNAs that interfere with Ras function
RT in Saccharomyces cerevisiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:2913-2917(1991).
RN [2]
RP PRELIMINARY SEQUENCE FROM N.A.
RX MEDLINE=95166216; PubMed=7862125;
RA Han L., Colicelli J.;
RT "A human protein selected for interference with Ras function
RT interacts directly with Ras and competes with Raf1.";
RL Mol. Cell. Biol. 15:1318-1323(1995).
RN [3]
RP SEQUENCE FROM N.A., AND REVISIONS.
RC TISSUE-Glioblastoma;
RA Han L., Colicelli J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP FUNCTION, DOMAINS, PHOSPHORYLATION, TISSUE SPECIFICITY, AND
RP ALTERNATIVE SPLICING (RINL-DELTA).
RX MEDLINE=97289700; PubMed=9144171;
RA Han L., Wong D., Dhaka A., Afar D., White M., Xie W., Herschman H.,
RA Witte O., Colicelli J.;
RT "Protein binding and signaling properties of RINL suggest a unique
RT effector function.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:4954-4959(1997).
CC -1- FUNCTION: MAY INTERFERE WITH RAS FUNCTION. INTERACTS DIRECTLY
CC WITH RAS AND COMPETES WITH RAF1 IN YEAST. FUNCTIONS AS AN EFFECTOR
CC OR REGULATOR OF RAS. MAY ALSO INTERACT WITH 14-3-3 PROTEINS AND
CC PROTEINS CONTAINING SH3 DOMAINS.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; RINL (SHOWN HERE) AND RINL-
CC DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED WITH HIGH
CC LEVELS IN BRAIN, PLACENTA AND PANCREAS.
CC -1- PTM: PROBABLY PHOSPHORYLATED.
CC -1- MISCELLANEOUS: RINL-DELTA SHOWS REDUCED ABILITY TO BIND TO RAS AND
CC 14-3-3 PROTEINS.


```

KW Ribosomal protein.
SQ SEQUENCE 105 AA; 11835 MW; 6C68DB611ED7F620 CRC64;

Query Match 80.6%; Score 29; DB 1; Length 105;
Best Local Similarity 71.4%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
   | | | | |
Db 33 TRGHLSK 39

RESULT 12
HIS3_RHOSH
ID HIS3_RHOSH STANDARD; PRT; 119 AA.
AC Q53158;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19) (PRA-CH).
GN HIS1.
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RC MEDLINE=96349112; PubMed=8760919;
RT Oriol E., Mendez-Alvarez S., Barbe J., Gibert I.;
RT "Cloning of the Rhodobacter sphaeroides hisL gene: unfunctionality
RT of the encoded protein and lack of linkage to other his genes.";
RT Microbiology 142:2071-2078(1996).
CC -!- CATALYTIC ACTIVITY: 1-(5-phosphoribosyl)-AMP + H(2)O - 1-(5-
CC phosphoribosyl)-5-[(3-
CC phosphoribosylamino)methylideneamino]imidazole-4-carboxamide.
CC -!- PATHWAY: Histidine biosynthesis; third step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE PRA-CH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X82010; CAAS7537.1; -.
CC InterPro; IPR002496; PRA-CH.
CC Pfam; PF01502; PRA-CH; 1.
CC ProDom; PD002610; PRA-CH; 1.
CC Histidine biosynthesis; Hydrolase.
CC SEQUENCE 119 AA; 13389 MW; 2682DFA9EC17B69F CRC64;

Query Match 80.6%; Score 29; DB 1; Length 119;
Best Local Similarity 71.4%; Pred. No. 16;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
   | | | | |
Db 64 TSGHVOR 70

RESULT 13
HEXP_LEIMA
ID HEXP_LEIMA STANDARD; PRT; 271 AA.
AC Q04832;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA-binding protein HEXPBP (Hexamer-binding protein).
GN HEXPBP.

```

OS Leishmania major.
 OC Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIH S;
 RX MEDLINE=93300785; PubMed=8314766;
 RA Webb J.R., McMaster W.R.;
 RT "Molecular cloning and expression of a Leishmania major gene encoding
 RT a single-stranded DNA-binding protein containing nine 'CCHC' zinc
 RT finger motifs";
 RL J. Biol. Chem. 268:13994-14002(1993).
 CC -1- FUNCTION: BINDS TO SINGLE-STRANDED DNA LOCATED IN THE 5'
 CC HEXANUCLEOTIDE REPEAT REGION OF THE L-MAJOR LEISHMANOLYSIN
 CC (GP63) GENE.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: CONTAINS 9 CCHC-TYPE ZINC FINGERS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M94390; AAA29245.1; -;
 DR PIR: S27822; S27822.
 DR InterPro: IPR001878; Znf_CCHC.
 DR Pfam: PF00098; Zf_CCHC; 9
 DR PRINTS: PRO00939; C2HCZNFINGER.
 DR SMART: SM00343; Znf_C2HC; 9.
 DR PROSITE: PS50158; Zf_CCHC; 9.
 KW DNA-binding; Zinc-finger; Nuclear protein; Repeat.
 FT ZN_FING 16 33 CCHC-TYPE 1.
 FT ZN_FING 43 60 CCHC-TYPE 2.
 FT ZN_FING 70 87 CCHC-TYPE 3.
 FT ZN_FING 97 114 CCHC-TYPE 4.
 FT ZN_FING 140 157 CCHC-TYPE 5.
 FT ZN_FING 168 185 CCHC-TYPE 6.
 FT ZN_FING 196 213 CCHC-TYPE 7.
 FT ZN_FING 222 239 CCHC-TYPE 8.
 FT ZN_FING 253 270 CCHC-TYPE 9.
 SQ SEQUENCE 271 AA; 28225 MW; ID33892EF9F743D8 CRC64;
 Query Match 80.6%; Score 29; DB 1; Length 271;
 Best Local Similarity 83.3%; Pred. No. 39;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SCHLSR 7
 Db 204 SCHMSR 209
 RESULT 14
 CNBP_MOUSE
 ID CNBP_MOUSE STANDARD; PRT; 170 AA.
 AC P53996;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cellular nucleic acid binding protein (CNBP).
 GN ZNF9 OR CNBP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=95203870; PubMed=7896269;
 RA Warden C.H., Krisans S.K., Purcell-Ruynh D., Leete L.M.,
 RA Daluiski A., Diep A., Taylor B.A., Lusis A.J.;

RT "Mouse cellular nucleic acid binding proteins: a highly conserved
 RT family identified by genetic mapping and sequencing.";
 RL Genomics 24:14-19(1994).
 CC -1- FUNCTION: SINGLE STRANDED DNA-BINDING PROTEIN, WITH SPECIFICITY TO
 CC THE STEROL REGULATORY ELEMENT (SRE). CNBP IS INVOLVED IN STEROL-
 CC MEDIATED REPRESSION.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ALSO PRESENT IN ENDOPLASMIC
 CC RETICULUM.
 CC -1- TISSUE SPECIFICITY: PRESENT IN ALL TISSUES EXAMINED.
 CC -1- SIMILARITY: TO S.POMBE BYR3 AND TO RETROVIRAL NUCLEIC ACID BINDING
 CC PROTEINS (NBP).
 CC -1- SIMILARITY: CONTAINS 7 CCHC-TYPE ZINC FINGERS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L12693; AAA89198.1; -;
 DR EMBL: Z11870; CAA77896.1; -;
 DR EMBL: X63866; CAA45345.1; -;
 DR EMBL: Z11871; CAA77897.1; -;
 DR EMBL: U20326; AAB60490.1; -;
 DR MGD: MGI:88431; Cnbp.
 DR InterPro: IPR001878; Znf_CCHC.
 DR Pfam: PF00098; Zf_CCHC; 7
 DR PRINTS: PRO00939; C2HCZNFINGER.
 DR SMART: SM00343; Znf_C2HC; 7.
 DR PROSITE: PS50158; Zf_CCHC; 7.
 KW Zinc-finger; DNA-binding; Transcription regulation; Repressor; Repeat.
 FT ZN_FING 4 21 CCHC-TYPE 1.
 FT ZN_FING 45 62 CCHC-TYPE 2.
 FT ZN_FING 65 82 CCHC-TYPE 3.
 FT ZN_FING 89 106 CCHC-TYPE 4.
 FT ZN_FING 110 127 CCHC-TYPE 5.
 FT ZN_FING 128 145 CCHC-TYPE 6.
 FT ZN_FING 149 166 CCHC-TYPE 7.
 FT VARIANT 2 2 S -> R (IN CLONE 6).
 FT VARIANT 33 33 G -> GGGFTSD (IN CLONE 4).
 FT VARIANT 65 65 D -> DE (IN CLONE 14).
 FT VARIANT 98 98 G -> D (IN CLONE 14).
 SQ SEQUENCE 170 AA; 18742 MW; 152BEC42881358E8 CRC64;
 Query Match 77.8%; Score 28; DB 1; Length 170;
 Best Local Similarity 83.3%; Pred. No. 39;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SCHLSR 7
 Db 157 SCHLAR 162
 RESULT 15
 CNBP_CHICK
 ID CNBP_CHICK STANDARD; PRT; 172 AA.
 AC Q42395;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cellular nucleic acid binding protein (CNBP).
 GN ZNF9 OR CNBP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98129058; PubMed=9467877;
 RA van Heumen W.R.A., Claxton C., Pickles J.O.;

RT "Sequence and tissue distribution of chicken cellular nucleic acid
RL binding protein cDNA."; 118B:659-665(1997).
CC Comp. Biochem. Physiol. 118B:659-665(1997).
CC -1- FUNCTION: SINGLE STRANDED DNA-BINDING PROTEIN, WITH SPECIFICITY TO
CC THE STEROL REGULATORY ELEMENT (SRE). CNBP IS INVOLVED IN STEROL-
CC MEDIATED REPRESSION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ALSO PRESENT IN ENDOPLASMIC
CC RETICULUM (BY SIMILARITY).
CC -1- SIMILARITY: TO S.POMBE BYR3 AND TO RETROVIRAL NUCLEIC ACID BINDING
CC PROTEINS (NBP).
CC -1- SIMILARITY: CONTAINS 7 CCHC-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF004942; AAB62243.1; -;
DR InterPro: IPR001878; Znf_CCHC.
DR Pfam: PF00098; zf-CCHC; 7.
DR PRINTS: PR00939; C2HCZNFINGER.
DR SMART: SM00343; Znf_C2HC; 7.
DR PROSITE: PS0158; ZF_CCHC; 7.
KW Zinc-finger; DNA-binding; Transcription regulation; Repressor; Repeat.
FT ZN_FING 4 21 CCHC-TYPE 1.
FT ZN_FING 46 63 CCHC-TYPE 2.
FT ZN_FING 67 84 CCHC-TYPE 3.
FT ZN_FING 91 108 CCHC-TYPE 4.
FT ZN_FING 112 129 CCHC-TYPE 5.
FT ZN_FING 130 147 CCHC-TYPE 6.
FT ZN_FING 151 168 CCHC-TYPE 7.
SQ SEQUENCE 172 AA; 19043 MW; 123F4E248A980390 CRC64;

Query Match 77.8%; Score 28; DB 1; Length 172;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGHLSR 7
| | | | |
Db 159 SGHLAR 164

Search completed: July 15, 2003, 12:53:30
Job time : 7.66667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:50:51 ; Search time 12.3333 Seconds
(without alignments)
54.563 Million cell updates/sec

Title: US-10-006-069A-247

Perfect score: 36

Sequence: 1 TSGHLSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	220	2 S45927	probable finger pr
2	33	91.7	231	2 S55078	hypothetical prote
3	32	88.9	214	2 T40349	vesicle transport
4	32	88.9	214	2 T42225	hypothetical prote
5	32	88.9	285	2 G85016	probable myb-relat
6	31	86.1	265	2 H96785	protein F10A5.22 [
7	31	86.1	286	2 T19449	hypothetical prote
8	31	86.1	357	2 A88633	protein F56B3.8 [1
9	31	86.1	423	1 W2VZ17	I7 protein - vacci
10	31	86.1	423	2 C42511	I7L protein - vacc
11	31	86.1	423	2 T37344	topoisomerase II -
12	31	86.1	496	2 C84678	hypothetical prote
13	31	86.1	503	2 JW0046	estrogen receptor
14	31	86.1	531	2 D70016	hypothetical prote
15	31	86.1	639	2 T47017	hypothetical prote
16	31	86.1	639	2 AH0236	probable membrane
17	31	86.1	762	2 T00410	protein kinase hom
18	30	83.3	400	2 T18570	probable malate de
19	30	83.3	680	2 G86672	Na+/H+ antiporter
20	30	83.3	783	1 A38637	Ras interactor RIN
21	30	83.3	1644	2 F91286	hypothetical prote
22	30	83.3	1644	2 B86128	hypothetical prote
23	30	83.3	1752	2 T48965	hypothetical prote
24	29	80.6	119	2 S49467	phosphoribosyl-AMP
25	29	80.6	135	2 S59649	hypothetical prote
26	29	80.6	271	2 A47156	hexamer-binding pr
27	29	80.6	367	2 H88035	protein M01D1.1 [1
28	29	80.6	389	2 E72537	probable cysteine
29	29	80.6	514	2 D71158	hypothetical prote

ALIGNMENTS

RESULT 1

S45927

probable finger protein YBR066c - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein YBR0616

C;Species: Saccharomyces cerevisiae

C;Date: 06-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 19-Apr-2002

C;Accession: S45927, S45926

R;Feldmann, H.; Mannhaupt, G.; Schwarze, C.; Vetter, I.

submitted to the Protein Sequence Database, August 1994

A;Reference number: S45927

A;Accession: S45927

A;Molecule type: DNA

A;Residues: 1-187 <FE>

A;Cross-references: EMBL:Z35935; MIPS:YBR066c

A;Experimental source: strain S288C

R;Domdey, H.; Gassenhuber, H.; Obermaier, B.; Piravandi, E.

submitted to the Protein Sequence Database, August 1994

A;Reference number: S45816

A;Accession: S45926

A;Molecule type: DNA

A;Residues: 124-220 <DOM>

A;Cross-references: EMBL:Z35935; MIPS:YBR066c

A;Experimental source: strain S288C

C;Genetics:

A;Gene: SGD:NRG2

A;Cross-references: SGD:S0000270

A;Map position: 2R

A;Note: YBR066c

C;Keywords: nucleus; zinc finger

F;155-175/Region: zinc finger

F;183-205/Region: zinc finger

Query Match 100.0%; Score 36; DB 2; Length 220;

Best Local Similarity 100.0%; Pred. No. 1.9;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TSGHLSR 7

Db 164 TSGHLSR 170

RESULT 2

S55078

hypothetical protein YDR043c - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein YD5112.01c

C;Species: Saccharomyces cerevisiae

C;Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002

C;Accession: S55078

R;Oliver, K.; Harris, D.

submitted to the EMBL Data Library, June 1995

A;Reference number: S55078

A;Accession: S55078

hypothetical colle
sensory transducti
hypothetical prote
hypothetical prote
hypothetical prote
probable regulator
hypothetical prote
hypothetical prote
hypothetical prote
universal minicirc
hypothetical prote
transcription regu
calcimedin, 32K -
cellular nucleic a
zinc finger protei
cellular nucleic a

A:Molecule type: DNA

A:Residues: 1-231 <OLI>
 A:Cross-references: EMBL:Z49812; NID:g1204147; PID:g854427; GSPDB:GN00004; MIPS:YDR043c
 A:Experimental source: strain AB972
 C:Genetics:
 A:Gene: SGD:NRG1; MIPS:YDR043c
 A:Cross-references: SGD:S0002450
 A:Map position: 4R

Query Match 91.7%; Score 33; DB 2; Length 231;
 Best Local Similarity 85.7%; Pred. No. 9.5;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
 |||||
 Db 185 TSGHLAR 191

RESULT 3

T40349
 vesicle transport v-snare protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T40349
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z21922
 A:Accession: T40349
 A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-214 <WO>
 A:Cross-references: EMBL:AL022070; PIDN:CAAL7790.1; GSPDB:GN000067; SPDB:SPBC3B9.10
 A:Experimental source: strain 972h-; cosmid c3B9
 C:Genetics:
 A:Gene: SPDB:SPBC3B9.10
 A:Map position: 2
 A:Introns: 9/2; 44/3; 81/3; 165/3; 186/2

Query Match 88.9%; Score 32; DB 2; Length 214;
 Best Local Similarity 85.7%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
 |||||
 Db 169 TSGHLDR 175

RESULT 4

T42225
 hypothetical protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
 C:Accession: T42225
 R:Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
 DNA Res. 4, 363-369, 1997
 A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
 A:Reference number: Z17323; MUID:98162722; PMID:9501991
 A:Accession: T42225
 A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-214 <YOS>
 A:Cross-references: EMBL:D89116; NID:g1749439; PIDN:BAAL13778.1; PID:g1749440
 A:Experimental source: strain PR745

Query Match 88.9%; Score 32; DB 2; Length 214;
 Best Local Similarity 85.7%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
 |||||
 Db 169 TSGHLDR 175

RESULT 5

G85016
 probable myb-related DNA-binding protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 C:Accession: G85016
 R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
 Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: A85001; MUID:20083488; PMID:10617198
 A:Accession: G85016
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-285 <STO>
 A:Cross-references: GB:NC_001268; NID:g7267625; PIDN:CAR80937.1; GSPDB:GN00140
 C:Genetics:

A:Gene: AT4G01280
 A:Map position: 4

Query Match 88.9%; Score 32; DB 2; Length 285;
 Best Local Similarity 85.7%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
 |||||
 Db 235 TSGHLQR 241

RESULT 6

H96785
 protein F10A5.22 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: H96785
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Aloni
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Mafti, R.; Marzia
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: H96785
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-265 <STO>
 A:Cross-references: GB:AB005173; NID:g9369368; PIDN:AAF87117.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F10A5.22
 A:Map position: 1

Query Match 86.1%; Score 31; DB 2; Length 265;
 Best Local Similarity 71.4%; Pred. No. 31;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
 |||||
 Db 172 TSGHAR 178

RESULT 7

T19449
 hypothetical protein C25D7.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
 C:Accession: T19449
 R:Ainscough, R.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19125
 A:Accession: T19449

A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-286 <WIL>
A:Cross-references: EMBL:Z81039; PIDN:CAB02773.1; GSPDB:GN00023; CESP:C25D7.2
A:Experimental source: clone C25D7
C:Genetics:

A:Gene: CESP:C25D7.2
A:Map position: 5

A:Introns: 28/3; 128/3; 212/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F36D3.4

Query Match 86.1%; Score 31; DB 2; Length 286;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGHLR 7
Db 55 SGHLR 60

RESULT 8

A88633 protein F56B3.8 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

A:Accession: A88633
R:anonymous, The C. elegans Sequencing Consortium.

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes

A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_elegans/

A:Accession: A88633
A>Status: preliminary
A:Molecule type: DNA

A:Residues: 1-357 <STO>
A:Cross-references: GB:chr.IV; PIDN:AAC02619.1; PID:G2854205; GSPDB:GN00022; CESP:F56B3.8

A:Note: similar to the L2P family of ribosomal proteins
C:Genetics:

A:Gene: F56B3.8
A:Map position: 4

Query Match 86.1%; Score 31; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
Db 168 TSGHLS 173

RESULT 9

WZV217

I7 protein - vaccinia virus (strain WR)

C:Species: vaccinia virus

C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jul-1999
A:Accession: G29889; A38497

R:Schmitt, J.F.C.; Stunnenberg, H.G.

J. Virol. 62, 1889-1897, 1988

A:Title: Sequence and transcriptional analysis of the vaccinia virus HindIII I fragment.

A:Reference number: A29889; MUID:88215015; PMID:2835495

A:Accession: G29889

A:Molecule type: DNA

A:Residues: 1-423 <SCH>

A:Cross-references: NID:G335662; PIDN:AAB59809.1; PID:G335669

R:Fathl, Z.; Condit, R.C.

Virology 181, 258-272, 1991

A:Title: Genetic and molecular biological characterization of a vaccinia virus temperature-sensitive mutant

A:Reference number: A38497; MUID:91134989; PMID:1994576

A:Accession: A38497

A:Molecule type: DNA

A:Residues: 1-215 <FAT>

A:Cross-references: EMBL:M29901

C:Genetics:

A:Gene: I7

C:Superfamily: vaccinia virus I7 protein

C:Keywords: late protein

Query Match 86.1%; Score 31; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGHLR 7
Db 74 SGHLR 79

RESULT 10

C42511

I7L protein - vaccinia virus (strain Copenhagen)

C:Species: vaccinia virus

A:Note: host Homo sapiens (man)

C:Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 08-Apr-1994

A:Accession: C42511

R:Johnson, G.P.

Submitted to GenBank, June 1990

A:Reference number: A33172

A:Accession: C42511

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-423 <JOH>

C:Superfamily: vaccinia virus I7 protein

Query Match 86.1%; Score 31; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGHLR 7
Db 74 SGHLR 79

RESULT 11

T37344

topoisomerase II - vaccinia virus (strain Ankara)

C:Species: vaccinia virus

A:Variety: strain Ankara

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

A:Accession: T37344

R:Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dorner, F.

Submitted to the EMBL Data Library, March 1997

A:Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain

A:Reference number: Z20877

A:Accession: T37344

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-423 <ANT>

A:Cross-references: EMBL:O94848; PIDN:AAB96439.1

A:Experimental source: strain Ankara

C:Genetics:

A:Note: MVA068L

C:Superfamily: vaccinia virus I7 protein

Query Match 86.1%; Score 31; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGHLR 7
Db 74 SGHLR 79

RESULT 12

C84678

hypothetical protein At2g27900 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84678
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84678
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-496 <STO>
A:Cross-references: GB:AE002093; NID:g4510426; PIDN:AAD21512.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g27900
A:Map position: 2

Query Match 86.1%; Score 31; DB 2; Length 496;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
|||||
DB 433 TSGHLS 438

RESULT 13

JW0046
estrogen receptor beta2 - rat
N:Alternate names: ERbeta2
C:Species: Rattus norvegicus (Norway rat)
C>Date: 17-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 17-Mar-1999
C:Accession: JW0046
R:Maruyama, K.; Endoh, H.; Sasaki-Iwaka, H.; Kanou, H.; Shimaya, E.; Hashimoto, S.; Kat
Biochem. Biophys. Res. Commun. 246, 142-147, 1998
A:Title: A novel isoform of rat estrogen receptor beta with 18 amino acid insertion in t
A:Reference number: JW0046; MUID:98262932; PMID:9600083
A:Accession: JW0046
A:Molecule type: mRNA
A:Residues: 1-503 <MAR>
A:Cross-references: DBJ:AB012721
C:Comment: This protein functions as a negative regulator of estrogen action.
C:Superfamily: estrogen receptor; erba transforming protein homology
F:102-381/Domain: erba transforming protein homology <ERBA>

Query Match 86.1%; Score 31; DB 2; Length 503;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
|||||
DB 37 TSGHLS 42

RESULT 14

D70016
hypothetical protein yun1 - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: D70016
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, K.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
A:Authors: Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, M

A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*
A:Reference number: A65580; MUID:98044033; PMID:9384377
A:Accession: D70016
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-531 <KUN>
A:Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAB15232.1; PID:el1843
A:Experimental source: strain 168
C:Genetics:
A:Gene: yun1

Query Match 86.1%; Score 31; DB 2; Length 531;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGHLSR 7
|||||
DB 417 SGHLSR 422

RESULT 15

T47017
hypothetical protein [imported] - Yersinia pestis
C:Species: Yersinia pestis
C>Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: T47017
R:Buchrieser, C.; Rusniok, C.; Couve, E.; Frangeul, L.; Billault, A.; Kunst, F.; Carn
submitted to the EMBL Data Library, October 1998
A:Description: DNA sequence of the 102 kbases unstable region of Yersinia pestis.
A:Reference number: Z24348
A:Accession: T47017
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-639 <BUC>
A:Cross-references: EMBL:AL031866; PIDN:CAA21360.1
A:Experimental source: strain 6/69

Query Match 86.1%; Score 31; DB 2; Length 639;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
|||||
DB 525 TSGHLS 530

Search completed: July 15, 2003, 12:55:32
Job time : 14.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:51:16 ; Search time 10.6667 Seconds
(without alignments)
19.309 Million cell updates/sec

Title: US-10-006-069A-68
Perfect score: 36
Sequence: 1 RSDHLR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pcp.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pcp.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pcp.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pcp.*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pcp.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	91.7	7	4 US-09-173-941-69	Sequence 69, Appl
2	33	91.7	21	4 US-09-229-007A-94	Sequence 94, Appl
3	33	91.7	21	4 US-09-229-007A-96	Sequence 96, Appl
4	33	91.7	26	2 US-08-620-151-3	Sequence 3, Appl
5	33	91.7	77	2 US-08-570-227A-5	Sequence 5, Appl
6	33	91.7	77	2 US-08-570-227A-6	Sequence 6, Appl
7	33	91.7	77	2 US-08-570-227A-7	Sequence 7, Appl
8	33	91.7	77	4 US-09-077-991-7	Sequence 7, Appl
9	33	91.7	77	4 US-09-077-991-8	Sequence 8, Appl
10	33	91.7	77	4 US-09-077-991-9	Sequence 9, Appl
11	33	91.7	94	4 US-09-229-007A-9	Sequence 9, Appl
12	33	91.7	98	4 US-09-229-007A-10	Sequence 10, Appl
13	33	91.7	168	2 US-08-353-476-73	Sequence 73, Appl
14	33	91.7	241	2 US-08-353-476-116	Sequence 116, App
15	33	91.7	273	2 US-08-353-476-112	Sequence 112, App
16	31	86.1	411	1 US-08-399-561-2	Sequence 2, Appl
17	30	83.3	77	2 US-08-570-227A-4	Sequence 4, Appl
18	30	83.3	77	2 US-08-570-227A-10	Sequence 10, Appl
19	30	83.3	77	4 US-09-077-991-6	Sequence 6, Appl
20	30	83.3	77	4 US-09-077-991-12	Sequence 12, Appl
21	30	83.3	288	4 US-09-492-985-2	Sequence 2, Appl
22	30	83.3	480	2 US-08-570-227A-2	Sequence 2, Appl
23	30	83.3	480	4 US-09-077-991-2	Sequence 2, Appl
24	29	80.6	302	2 US-08-203-532F-4	Sequence 4, Appl
25	29	80.6	302	3 US-08-950-860-16	Sequence 16, Appl
26	29	80.6	302	4 US-09-078-465-4	Sequence 4, Appl
27	29	80.6	302	5 PCT-US95-01882A-4	Sequence 4, Appl

28	29	80.6	303	2	US-08-203-532F-2	Sequence 2, Appl
29	29	80.6	303	4	US-09-078-465-2	Sequence 2, Appl
30	29	80.6	303	5	PCT-US95-01882A-2	Sequence 2, Appl
31	29	80.6	667	2	US-08-718-661-2	Sequence 2, Appl
32	28	77.8	7	4	US-09-173-941-71	Sequence 71, Appl
33	28	77.8	10	3	US-09-139-762A-19	Sequence 19, Appl
34	28	77.8	10	3	US-09-139-762A-53	Sequence 53, Appl
35	28	77.8	20	1	US-08-040-548-19	Sequence 19, Appl
36	28	77.8	20	1	US-08-466-344-19	Sequence 19, Appl
37	28	77.8	21	4	US-09-229-007A-95	Sequence 95, Appl
38	28	77.8	24	4	US-08-874-569B-16	Sequence 16, Appl
39	28	77.8	26	2	US-08-620-151-111	Sequence 111, App
40	28	77.8	28	1	US-08-040-548-34	Sequence 34, Appl
41	28	77.8	28	1	US-08-466-344-34	Sequence 34, Appl
42	28	77.8	28	4	US-09-058-459-1	Sequence 1, Appl
43	28	77.8	28	4	US-09-127-926-1	Sequence 1, Appl
44	28	77.8	28	4	US-09-037-179B-15	Sequence 15, Appl
45	28	77.8	28	4	US-09-240-179-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-173-941-69
; Sequence 69, Application US/09173941
; Patent No. 6140081
; GENERAL INFORMATION:
; APPLICANT: BARBAS, Carlos F.
; TITLE OF INVENTION: ZINC FINGER BINDING DOMAINS FOR GNN
; FILE REFERENCE: NOV00815
; CURRENT APPLICATION NUMBER: US/09/173,941
; CURRENT FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 69
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: nucleotide
; OTHER INFORMATION: codon binding sequence
US-09-173-941-69

Query Match	91.7%	Score 33;	DB 4;	Length 7;
Best Local Similarity	85.7%	Pred No. 2e+05;		
Matches	6;	Conservative	1;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	RSDHLR 7		
Db	1	RSDHLTR 7		

RESULT 2
US-09-229-007A-94
; Sequence 94, Application US/09229007A
; Patent No. 6453242
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Case, Casey C.
; APPLICANT: Cox III, George N.
; APPLICANT: Jamieson, Andrew
; APPLICANT: Rebar, Edward J.
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Selection of Sites for Targeting by Zinc Finger
; TITLE OF INVENTION: Proteins and Methods of Designing Zinc Finger Proteins
; FILE REFERENCE: to Bind to Preselcted Sites
; FILE REFERENCE: 019496-001800US
; CURRENT APPLICATION NUMBER: US/09/229,007A
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 94

```
/
/ LENGTH: 21
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:ZFP sequence
/ OTHER INFORMATION: (F1, F2 and F3) from SBS design GR-223
US-09-229-007A-94

Query Match          91.7%; Score 33; DB 4; Length 21;
Best Local Similarity 85.7%; Pred. No. 0.95;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 8 RSDHLTR 14

RESULT 3
US-09-229-007A-96
; Sequence 96, Application US/09229007A
; Patent No. 6453242
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Case, Casey C.
; APPLICANT: Cox III, George N.
; APPLICANT: Jamieson, Andrew
; APPLICANT: Rebar, Edward J.
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Selection of Sites for Targeting by Zinc Finger
; TITLE OF INVENTION: Proteins and Methods of Designing Zinc Finger Proteins
; TITLE OF INVENTION: to Bind to Preslected Sites
; FILE REFERENCE: 019496-001800US
; CURRENT APPLICATION NUMBER: US/09/229, 007A
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:ZFP sequence
; OTHER INFORMATION: (F1, F2, F3) from SP1
US-09-229-007A-96

Query Match          91.7%; Score 33; DB 4; Length 21;
Best Local Similarity 85.7%; Pred. No. 0.95;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 15 RSDHLSK 21

RESULT 4
US-08-620-151-3
; Sequence 3, Application US/08620151
; Patent No. 5928955
; GENERAL INFORMATION:
; APPLICANT: Imperiall, Barbara
; APPLICANT: Walkup, Grant K.
; TITLE OF INVENTION: PEPTIDYL FLUORESCENT CHEMOSENSOR FOR
; TITLE OF INVENTION: DIVALENT ZINC
; NUMBER OF SEQUENCES: 136
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
```

```
/
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/620,151
/ FILING DATE: 22-MAR-1996
/ CLASSIFICATION: 422
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Shannon, Karen L.
/ REGISTRATION NUMBER: 36,675
/ REFERENCE/DOCKET NUMBER: 8597/6
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312-321-4200
/ TELEFAX: 312-321-4299
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 26 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-08-620-151-3

Query Match          91.7%; Score 33; DB 2; Length 26;
Best Local Similarity 85.7%; Pred. No. 1.2;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 13 RSDHLSK 19

RESULT 5
US-08-570-227A-5
; Sequence 5, Application US/08570227A
; Patent No. 5981217
; GENERAL INFORMATION:
; APPLICANT: Subramaniam, M.
; APPLICANT: Spelsberg, T. C.
; TITLE OF INVENTION: DNA ENCODING TGF-BETA INDUCIBLE
; TITLE OF INVENTION: EARLY FACTOR-1 (TIEF-1), A GENE EXPRESSED
; TITLE OF INVENTION: BY OSTEOBLASTS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,227A
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D.
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.157US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
```

;
; LENGTH: 77 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
US-08-570-227A-5

Query Match 91.7%; Score 33; DB 2; Length 77;
Best Local Similarity 85.7%; Pred. No. 3.8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
| | | | |
Db 70 RSDHLSK 76

RESULT 6

US-08-570-227A-6
; Sequence 6, Application US/08570227A
; Patent No. 5981217

GENERAL INFORMATION:

; APPLICANT: Subramaniam, M.
; APPLICANT: Spelsberg, T. C.
; TITLE OF INVENTION: DNA ENCODING TGF-BETA INDUCIBLE
; TITLE OF INVENTION: EARLY FACTOR-1 (TIEF-1), A GENE EXPRESSED
; TITLE OF INVENTION: BY OSTEOBLASTS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,227A
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 150.1570S1

TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-359-3260
TELEFAX: 612-359-3263
TELEX:

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 77 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-570-227A-6

Query Match 91.7%; Score 33; DB 2; Length 77;
Best Local Similarity 85.7%; Pred. No. 3.8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
| | | | |
Db 70 RSDHLSK 76

RESULT 7

US-08-570-227A-7
; Sequence 7, Application US/08570227A
; Patent No. 5981217

GENERAL INFORMATION:

; APPLICANT: Subramaniam, M.
; APPLICANT: Spelsberg, T. C.
; TITLE OF INVENTION: DNA ENCODING TGF-BETA INDUCIBLE
; TITLE OF INVENTION: EARLY FACTOR-1 (TIEF-1), A GENE EXPRESSED
; TITLE OF INVENTION: BY OSTEOBLASTS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/570,227A

FILING DATE: 11-DEC-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Woessner, Warren D

REGISTRATION NUMBER: 30,440

REFERENCE/DOCKET NUMBER: 150.1570S1

TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-359-3260
TELEFAX: 612-359-3263
TELEX:

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 77 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-570-227A-7

Query Match 91.7%; Score 33; DB 2; Length 77;
Best Local Similarity 85.7%; Pred. No. 3.8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
| | | | |
Db 70 RSDHLSK 76

RESULT 8

US-09-077-991-7

; Sequence 7, Application US/09077991

; Patent No. 6207375

GENERAL INFORMATION:

; APPLICANT: Subramaniam, M.

; APPLICANT: Spelsberg, T. C.

; APPLICANT: Roche, P. C.

; TITLE OF INVENTION: TGF-Beta inducible early factor-1

; TITLE OF INVENTION: (TIEF-1) and a method to detect breast cancer

; FILE REFERENCE: 150.1570S2

; CURRENT APPLICATION NUMBER: US/09/077,991

; CURRENT FILING DATE: 1998-07-07

; EARLIER APPLICATION NUMBER: PCT/US96/19555

; EARLIER FILING DATE: 1996-12-11

; EARLIER APPLICATION NUMBER: US 08/570,227

; EARLIER FILING DATE: 1995-12-11

; NUMBER OF SEQ ID NOS: 13

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-077-991-7

Query Match          91.7%; Score 33; DB 4; Length 77;
Best Local Similarity 85.7%; Pred. No. 3.8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 70 RSDHLSK 76

RESULT 9
US-09-077-991-8
; Sequence 8, Application US/09077991
; Patent No. 6207375
; GENERAL INFORMATION:
; APPLICANT: Subramaniam, M.
; APPLICANT: Spelsberg, T.C.
; APPLICANT: Roche, P.C.
; TITLE OF INVENTION: TGF-Beta Inducible early factor-1
; FILE REFERENCE: (TIEF-1) and a method to detect breast cancer
; FILE REFERENCE: 150.157US2
; CURRENT APPLICATION NUMBER: US/09/077,991
; CURRENT FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: PCT/US96/19555
; EARLIER FILING DATE: 1996-12-11
; EARLIER APPLICATION NUMBER: US 08/570,227
; EARLIER FILING DATE: 1995-12-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-077-991-8

Query Match          91.7%; Score 33; DB 4; Length 77;
Best Local Similarity 85.7%; Pred. No. 3.8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 70 RSDHLSK 76

RESULT 10
US-09-077-991-9
; Sequence 9, Application US/09077991
; Patent No. 6207375
; GENERAL INFORMATION:
; APPLICANT: Subramaniam, M.
; APPLICANT: Spelsberg, T.C.
; APPLICANT: Roche, P.C.
; TITLE OF INVENTION: TGF-Beta Inducible early factor-1
; FILE REFERENCE: (TIEF-1) and a method to detect breast cancer
; FILE REFERENCE: 150.157US2
; CURRENT APPLICATION NUMBER: US/09/077,991
; CURRENT FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: PCT/US96/19555
; EARLIER FILING DATE: 1996-12-11
; EARLIER APPLICATION NUMBER: US 08/570,227
; EARLIER FILING DATE: 1995-12-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-077-991-9

Query Match          91.7%; Score 33; DB 4; Length 77;
Best Local Similarity 85.7%; Pred. No. 3.8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 70 RSDHLSK 76

RESULT 11
US-09-229-007A-9
; Sequence 9, Application US/09229007A
; Patent No. 6453242
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Case, Casey C.
; APPLICANT: Cox III, George N.
; APPLICANT: Jamieson, Andrew
; APPLICANT: Rebar, Edward J.
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Selection of Sites for Targeting by Zinc Finger
; TITLE OF INVENTION: Proteins and Methods of Designing Zinc Finger Proteins
; FILE REFERENCE: 019496-001800US
; CURRENT APPLICATION NUMBER: US/09/229,007A
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:amino acids
; OTHER INFORMATION: 531-624 in Sp-1 transcription factor
US-09-229-007A-9

Query Match          91.7%; Score 33; DB 4; Length 94;
Best Local Similarity 85.7%; Pred. No. 4.7;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 78 RSDHLSK 84

RESULT 12
US-09-229-007A-10
; Sequence 10, Application US/09229007A
; Patent No. 6453242
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Case, Casey C.
; APPLICANT: Cox III, George N.
; APPLICANT: Jamieson, Andrew
; APPLICANT: Rebar, Edward J.
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Selection of Sites for Targeting by Zinc Finger
; TITLE OF INVENTION: Proteins and Methods of Designing Zinc Finger Proteins
; FILE REFERENCE: 019496-001800US
; CURRENT APPLICATION NUMBER: US/09/229,007A
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Sp-1
```

; OTHER INFORMATION: transcription factor consensus sequence
US-09-229-007A-10

Query Match 91.7%; Score 33; DB 4; Length 98;
Best Local Similarity 85.7%; Pred. No. 4.9;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSDHLSR 7
Db 84 RSDHLSK 90

RESULT 13

US-08-353-476-73
; Sequence 73, Application US/08353476

; Patent No. 5871902

; GENERAL INFORMATION:

; APPLICANT: Weininger, Susan

; APPLICANT: Weininger, Arthur M

; TITLE OF INVENTION: METHOD OF DETECTION OF DNA WITH A

; TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION

; NUMBER OF SEQUENCES: 117

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Saliwanchik & Saliwanchik

; STREET: 2421 N.W. 41st St., Suite A-1

; CITY: Gainesville

; STATE: Florida

; COUNTRY: USA

; ZIP: 32606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/353,476

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Bencen, Gerard H

; REGISTRATION NUMBER: 35,746

; REFERENCE/DOCKET NUMBER: GP-100

; TELEPHONE: (904) 375-8100

; TELEFAX: (904) 372-5800

; INFORMATION FOR SEQ ID NO: 73:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 168 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: internal

; US-08-353-476-73

Query Match 91.7%; Score 33; DB 2; Length 168;

Best Local Similarity 85.7%; Pred. No. 8.7;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSDHLSR 7
Db 80 RSDHLSK 86

RESULT 14

US-08-353-476-116

; Sequence 116, Application US/08353476

; Patent No. 5871902

; GENERAL INFORMATION:

; APPLICANT: Weininger, Susan

; APPLICANT: Weininger, Arthur M

; TITLE OF INVENTION: METHOD OF DETECTION OF DNA WITH A

; TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION

; NUMBER OF SEQUENCES: 117

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Saliwanchik & Saliwanchik

; STREET: 2421 N.W. 41st St., Suite A-1

; CITY: Gainesville

; STATE: Florida

; COUNTRY: USA

; ZIP: 32606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/353,476

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Bencen, Gerard H

; REGISTRATION NUMBER: 35,746

; REFERENCE/DOCKET NUMBER: GP-100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (904) 375-8100

; TELEFAX: (904) 372-5800

; INFORMATION FOR SEQ ID NO: 116:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 241 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

; US-08-353-476-116

Query Match 91.7%; Score 33; DB 2; Length 241;

Best Local Similarity 85.7%; Pred. No. 13;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSDHLSR 7

Db 153 RSDHLSK 159

RESULT 15

US-08-353-476-112

; Sequence 112, Application US/08353476

; Patent No. 5871902

; GENERAL INFORMATION:

; APPLICANT: Weininger, Susan

; APPLICANT: Weininger, Arthur M

; TITLE OF INVENTION: METHOD OF DETECTION OF DNA WITH A

; TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION

; NUMBER OF SEQUENCES: 117

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Saliwanchik & Saliwanchik

; STREET: 2421 N.W. 41st St., Suite A-1

; CITY: Gainesville

; STATE: Florida

; COUNTRY: USA

; ZIP: 32606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/353,476

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Bencen, Gerard H

; REGISTRATION NUMBER: 35,746
; REFERENCE/DOCKET NUMBER: GP-100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-353-476-112

Query Match 91.7%; Score 33; DB 2; Length 273;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 RSDHLSR 7
Db 185 RSDHLSK 191

Search completed: July 15, 2003, 12:56:09
Job time : 11.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:50:16 ; Search time 23.6667 Seconds
(without alignments)
60.944 Million cell updates/sec

Title: US-10-006-069a-68
Perfect score: 36
Sequence: 1 RSDHLSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	1363	5 Q20766	Q20766 caenorhabdi
2	33	91.7	237	6 P79289	P79289 sus scrofa
3	33	91.7	278	13 Q800U4	Q800U4 xenopus lae
4	33	91.7	347	5 Q95R58	Q95R58 drosophila
5	33	91.7	415	11 Q9WTQ3	Q9WTQ3 rattus norv
6	33	91.7	426	5 Q9VPQ5	Q9VPQ5 drosophila
7	33	91.7	428	11 Q8V167	Q8V167 mus musculu
8	33	91.7	431	4 Q8TDD2	Q8TDD2 homo sapien
9	33	91.7	452	4 Q96MJ1	Q96MJ1 homo sapien
10	33	91.7	467	11 Q64167	Q64167 mus musculu
11	33	91.7	760	13 Q90WR9	Q90WR9 gallus gall
12	33	91.7	781	11 Q89090	Q89090 mus musculu
13	33	91.7	782	11 Q62445	Q62445 mus musculu
14	33	91.7	783	11 Q63158	Q63158 rattus norv
15	33	91.7	784	11 Q89087	Q89087 mus musculu
16	33	91.7	968	5 Q9VCB2	Q9VCB2 drosophila

17	33	91.7	1464	11 Q60488	Q60488 cavia porce
18	32	88.9	824	10 Q8RUF4	Q8RUF4 oryza sativ
19	32	88.9	972	5 Q9VLX0	Q9VLX0 drosophila
20	32	88.9	1332	10 Q9FJ71	Q9FJ71 arabidopsis
21	31	86.1	98	2 Q05687	Q05687 mycobacteri
22	31	86.1	168	5 Q9N492	Q9N492 caenorhabdi
23	31	86.1	195	11 Q9CZY8	Q9CZY8 mus musculu
24	31	86.1	204	11 Q9CVL9	Q9CVL9 mus musculu
25	31	86.1	212	11 Q9CRJ0	Q9CRJ0 mus musculu
26	31	86.1	300	5 Q9BHE3	Q9BHE3 leishmania
27	31	86.1	309	5 P91329	P91329 caenorhabdi
28	31	86.1	361	10 Q9SGA2	Q9SGA2 arabidopsis
29	31	86.1	410	5 Q9W1W2	Q9W1W2 drosophila
30	31	86.1	497	5 Q9T264	Q9T264 caenorhabdi
31	31	86.1	543	5 Q46016	Q46016 caenorhabdi
32	31	86.1	546	13 Q93567	Q93567 gallus gall
33	31	86.1	565	11 Q88939	Q88939 mus musculu
34	31	86.1	569	11 Q9QZ48	Q9QZ48 rattus norv
35	31	86.1	584	4 Q95365	Q95365 homo sapien
36	31	86.1	619	4 Q73453	Q73453 homo sapien
37	31	86.1	619	11 Q8VCZ7	Q8VCZ7 mus musculu
38	31	86.1	746	11 Q9D9X5	Q9D9X5 mus musculu
39	31	86.1	750	5 Q01348	Q01348 drosophila
40	31	86.1	801	5 Q9VTD3	Q9VTD3 drosophila
41	31	86.1	1210	3 Q8TFX9	Q8TFX9 aspergillus
42	30	83.3	100	5 Q77251	Q77251 drosophila
43	30	83.3	120	11 Q9DCC0	Q9DCC0 mus musculu
44	30	83.3	172	4 Q9UK33	Q9UK33 homo sapien
45	30	83.3	172	11 Q9DB38	Q9DB38 mus musculu

ALIGNMENTS

RESULT 1

Q20766 Q20766 PRELIMINARY; PRT; 1363 AA.
AC Q20766;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE F34D1.5 protein.
GN F34D1.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lennard N.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z77132; CAB00861.2; -.
DR InterPro; IPR002111; Cat_channel_TrpL.
DR InterPro; IPR000636; M-channel_nlg.
DR Pfam; PF00520; ion.trans; 2.
SQ SEQUENCE 1363 AA; 156770 MW; DFE8960976A4E0B9 CRC64;

Query Match 100.0%; Score 36; DB 5; Length 1363;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSDHLSR 7
| | | | |
Db 14 RSDHLSR 20

RESULT 2

P79289 ID P79289 PRELIMINARY; PRT; 237 AA.
 AC Q8UUU4
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE SPl transcription factor (Fragment).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ENDOMETRIUM;
 RX MEDLINE=20150569; PubMed=10687861;
 RA Simmen R.C.M., Zhang X.L., Zhang D., Wang Y., Michel F.J.,
 RA Simmen F.A.;
 RT "Expression and regulatory function of the transcription factor SPl in
 RT the uterine endometrium at early pregnancy: implications for
 RT epithelial phenotype.";
 RL Mol. Cell. Endocrinol. 159:159-170(2000).
 DR EMBL; U57347; AAB39513.3; -;
 DR HSSP; P08047; 1SP2.
 DR InterPro; IPR000822; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 3.
 DR ProDom; PD000003; Znf_C2H2; 2.
 DR SMART; SM00355; Znf_C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.
 KW DNA-binding; Metal-binding; Zinc-finger.
 FT NON_TER 1
 FT NON_TER 237 237
 SQ SEQUENCE 237 AA; 25421 MW; C6950DB42912DAB6 CRC64;

Query Match 91.7%; Score 33; DB 6; Length 237;
 Best Local Similarity 85.7%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RSDHLSR 7
 Db 167 RSDHLSK 173

RESULT 3

Q8UUU4 ID Q8UUU4 PRELIMINARY; PRT; 278 AA.
 AC Q8UUU4
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE SPl-like zinc-finger protein XSPR-2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ossipova O., Stick R., Pieler T.;
 RT "Interaction between two novel SPl-like zinc finger proteins and
 RT Brachyury in Xenopus.";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY062263; AAL47216.1; -;
 DR InterPro; IPR000822; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 3.
 DR PRINTS; ZN00048; ZINC_FINGER.
 DR ProDom; PD000003; Znf_C2H2; 1.
 DR SMART; SM00355; Znf_C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
 KW ZINC.
 SQ SEQUENCE 278 AA; 31098 MW; 44679B6E94E516D CRC64;

Query Match 91.7%; Score 33; DB 13; Length 278;
 Best Local Similarity 85.7%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RSDHLSR 7
 Db 249 RSDHLSK 255

RESULT 4
 Q95R58 ID Q95R58 PRELIMINARY; PRT; 347 AA.
 AC Q95R58;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE SD05726p.
 GN EP2237 OR CG4427.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Ceiniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY061602; AAL29150.1; -;
 DR FlyBase; FBgn0043364; EP2237.
 DR InterPro; IPR000822; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 3.
 DR ProDom; PD000003; Znf_C2H2; 2.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
 KW DNA-binding; Zinc-finger.
 SQ SEQUENCE 347 AA; 37832 MW; 153C660EEB49DDEF CRC64;

Query Match 91.7%; Score 33; DB 5; Length 347;
 Best Local Similarity 85.7%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RSDHLSR 7
 Db 253 RSDHLSK 259

RESULT 5
 Q9WTQ3 ID Q9WTQ3 PRELIMINARY; PRT; 415 AA.
 AC Q9WTQ3;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Kruppel-like transcription factor.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Uchida S.;
 RT "Kidney-enriched Kruppel-like transcription factor.";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB020759; BAA78378.1; -;
 DR HSSP; P08047; 1SP2.
 DR InterPro; IPR000822; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 3.
 DR ProDom; PD000003; Znf_C2H2; 1.


```

DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 415 AA; 44147 MW; 9569CC3ACE98722F CRC64;

Query Match          91.7%; Score 33; DB 11; Length 415;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 391 RSDHLSK 397

RESULT 6
Q9VPQ5 Q9VPQ5 PRELIMINARY; PRT; 426 AA.
AC Q9VPQ5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG4427 protein.
GN EP2237 OR CG4427.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.F., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Garg N.S., Geibart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003589; AAF51489.1; -.
DR HSSP; P08047; 1SP2.

```

```

DR FlyBase; FBgn0043364; EP2237.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 426 AA; 46491 MW; 7D3626F2C2CA0AE1 CRC64;

Query Match          91.7%; Score 33; DB 5; Length 426;
Best Local Similarity 85.7%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 332 RSDHLSK 338

RESULT 7
Q8VI67 Q8VI67 PRELIMINARY; PRT; 428 AA.
AC Q8VI67;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Osterix.
GN OSX OR C22.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21652532; PubMed=11792318;
RA Nakashima K., Zhou X., Kunkel G., Zhang Z., Deng J.M., Behringer R.R.,
RA de Crombrughe B.;
RT "The Novel Zinc Finger-Containing Transcription Factor Osterix is
RL Required for Osteoblast Differentiation and Bone Formation.";
RL Cell 108:17-29(2002).
DR EMBL; AF184902; AAL60067.1; -.
DR MGD; MGI:2159409; Osx.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
SQ SEQUENCE 428 AA; 44718 MW; B79498958743586 CRC64;

Query Match          91.7%; Score 33; DB 11; Length 428;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 362 RSDHLSK 368

RESULT 8
Q8TDD2 Q8TDD2 PRELIMINARY; PRT; 431 AA.
AC Q8TDD2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Osterix.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=OSTEOSARCOMA;
RA Ganss B.W.;
RT "cDNA sequence, gene structure and chromosomal localization of the
RT human osterix (OSX) gene.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF477981; AAL84281.1; -.
SQ SEQUENCE 431 AA; 44994 MW; 454A6FEA84309FF9 CRC64;

Query Match 91.7%; Score 33; DB 4; Length 431;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 365 RSDHLSK 371

RESULT 9
Q96MJ1 PRELIMINARY; PRT; 452 AA.
AC Q96MJ1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CDNA FLJ32295 fis, clone PROST2001823, weakly similar to transcription
DE factor SPL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotata T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J., Chiba Y.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Ishii S.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Iseno Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK056857; BAB71297.1; -.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF00096; zf-C2H2; 3.
DR ProDom: PD000003; Znf_C2H2; 2.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 3.
KW DNA-binding; Zinc-finger.
SQ SEQUENCE 452 AA; 45836 MW; 95383C60C112320F CRC64;

Query Match 91.7%; Score 33; DB 4; Length 452;
Best Local Similarity 85.7%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 389 RSDHLSK 395

RESULT 10
Q64167 PRELIMINARY; PRT; 467 AA.
AC Q64167; Q62251;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Trans-acting transcription factor 1 (Transcription factor SPL) (SPL
DE gene) (3' end).
GN SPL.
OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96016118; PubMed=7568082;
RA Persengiev S.P., Saffer J.D., Kilpatrick D.L.;
RT "An alternatively spliced form of the transcription factor Spl
RT containing only a single glutamine-rich transactivation domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:9107-9111(1995).
RN [2]
RP SEQUENCE OF 367-467 FROM N.A.
RX MEDLINE=92338398; PubMed=1633330;
RA Chestier A., Charnay P.;
RT "Difference in the genomic organizations of the related transcription
RT factors Spl and Krox-20; possible evolutionary significance.";
RL DNA Seq. 2:325-327(1992).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; S79832; AAB35321.1; -.
DR EMBL; X60136; CAA42721.1; -.
DR HSSP; P08047; 1SPI.
DR MGD; MGI:98372; Spl.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF00096; zf-C2H2; 3.
DR PRINTS; PR00048; ZINC_FINGER.
DR ProDom: PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 3.
KW DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ SEQUENCE 467 AA; 48749 MW; 4A373D67E6128197 CRC64;

Query Match 91.7%; Score 33; DB 11; Length 467;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 381 RSDHLSK 387

RESULT 11
Q90WR9 PRELIMINARY; PRT; 760 AA.
ID Q90WR9;
AC Q90WR9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Transcription factor.
GN SPL.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Chamboredon S., Hurault J., Vial E., Oliviero S., Castellazzi M.;
RT "Transcriptional control of SPARC by APl: role of Spl/3.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ317960; CAC84904.1; -.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF00096; zf-C2H2; 3.
DR ProDom: PD000003; Znf_C2H2; 2.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 3.
KW DNA-binding; Zinc-finger.
SQ SEQUENCE 760 AA; 78561 MW; 4ECBBE9DB49A8594 CRC64;

Query Match 91.7%; Score 33; DB 13; Length 760;
Best Local Similarity 85.7%; Pred. No. 71;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```
QY 1 RSDHLSR 7
DB 670 RSDHLSK 676

RESULT 12
ID O62445 PRELIMINARY; PRT; 781 AA.
AC O62445
DT 01-NOV-1996 (TREMblrel. 08, Created)
DT 01-NOV-1996 (TREMblrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Transcription factor Sp4.
GN SP4 OR SP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-NEUROBLASTOMA;
RX MEDLINE-96290594; PubMed-9628590;
RA Iyama S., Lee S.H., Minowa T., Mouradian M.M.;
RT "Sp family transcription factors regulate expression of rat D2
RT dopamine receptor gene.";
RL DNA Cell Biol. 17:471-479(1998).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF062566; AAC16484.1; -
DR HSSP; P08047; ISPI.
DR MGD; MGI:107372; Sp4.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
KW DNA-binding; Metal-binding; Nuclear protein; zinc-finger.
SQ SEQUENCE 781 AA; 80486 MW; 14CD12B8C58CF921 CRC64;

Query Match 91.7%; Score 33; DB 11; Length 781;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
DB 695 RSDHLSK 701

RESULT 13
ID O62445 PRELIMINARY; PRT; 782 AA.
AC O62445
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Zinc finger transcription factor Sp4.
GN SP4 OR SP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-NEUROBLASTOMA;
RX MEDLINE-96290594; PubMed-9628590;
RA Supp D.M., Witte D.P., Branford W.W., Smith E.P., Potter S.S.;
RT "Sp4, a member of the Sp1-family of zinc finger transcription factors,
RT is required for normal murine growth, viability, and male fertility.";
RL Dev. Biol. 176:284-299(1996).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; U62522; AAC52653.1; -
DR HSSP; P08047; ISPI.
DR TRANSFAC; T02414; -
```

```
DR MGD; MGI:107595; Sp4.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
KW DNA-binding; Metal-binding; Nuclear protein; zinc-finger.
SQ SEQUENCE 782 AA; 81966 MW; 50BDA3D6C0848A4F CRC64;

Query Match 91.7%; Score 33; DB 11; Length 782;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
DB 716 RSDHLSK 722

RESULT 14
ID O63158 PRELIMINARY; PRT; 783 AA.
AC O63158
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Zinc finger protein.
GN HF-1B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-HEART;
RX MEDLINE-93309478; PubMed-8321243;
RA Zhu H., Nguyen V.T.B., Brown A.B., Pourhossaini A., Garcia A.V.,
RA van Bilsen M., Chien K.R.;
RT "A Novel, Tissue-Restricted Zinc Finger Protein (HF-1b) Binds to the
RT Cardiac Regulatory Element (HF-1b/MEF-2) in the Rat Myosin Light
RT Chain-2 Gene.";
RL Mol. Cell. Biol. 13:4432-4444(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-HEART;
RA Zhu H.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; U07610; AAA17375.1; -
DR HSSP; P08047; ISPI.
DR TRANSFAC; T02456; -
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
KW DNA-binding; Metal-binding; Nuclear protein; zinc-finger.
SQ SEQUENCE 783 AA; 81984 MW; 394AA6D4DEA5A197 CRC64;

Query Match 91.7%; Score 33; DB 11; Length 783;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
DB 717 RSDHLSK 723

RESULT 15
O89087
```

```

ID O89087 . PRELIMINARY; PRT; 784 AA.
AC O89087;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Transcription factor SP1.
GN SP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=THYMUS;
RA Park E.J., Kim J.H., Kim C.G., Park S.D., Hong S.S.;
RT "Isolation of cDNA encoding transcription factor spl containing two
RT glutamine-rich transactivation domain.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF022363; AAC08527.1; -.
DR HSSP; P08047; 1spl.
DR MGD; MGI:98372; Spl.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF000096; zf-C2H2; 3.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ SEQUENCE 784 AA; 80705 MW; F658BEF0A64E146A CRC64;

Query Match 91.7%; Score 33; DB 11; Length 784;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLR 7
DB 698 RSDLSK 704

```

Search completed: July 15, 2003, 12:54:49
 Job time : 25.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:49:31 ; Search time 6.66667 seconds
(without alignments)
43.550 Million cell updates/sec

Title: US-10-006-069a-68

Perfect score: 36

Sequence: 1 RSDHLSR 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	33	91.7	415	1 KLF_MOUSE	Q9epw2 mus musculus
2	33	91.7	416	1 KLF_MOUSE	Q9u1h9 homo sapien
3	33	91.7	644	1 BTD_HUMAN	Q24266 drosophila
4	33	91.7	784	1 SP4_HUMAN	Q02446 homo sapien
5	33	91.7	785	1 SP1_HUMAN	P08047 homo sapien
6	33	91.7	788	1 SP1_RAT	Q01714 rattus norv
7	31	86.1	359	1 KLF8_HUMAN	Q95600 homo sapien
8	31	86.1	411	1 LCB_MOUSE	Q95600 homo sapien
9	30	83.3	152	1 SP6_MOUSE	Q95276 synchococc
10	30	83.3	244	1 BTE1_HUMAN	Q95276 mus musculus
11	30	83.3	244	1 BTE1_MOUSE	Q13886 homo sapien
12	30	83.3	244	1 BTE1_MOUSE	Q35739 mus musculus
13	30	83.3	244	1 BTE1_RAT	Q01713 rattus norv
14	30	83.3	251	1 BTE4_MOUSE	P58334 mus musculus
15	30	83.3	252	1 BTE4_HUMAN	Q9bxx1 homo sapien
16	30	83.3	288	1 KLF_MOUSE	Q9y2y9 h krueppel
17	30	83.3	289	1 KLF_MOUSE	Q9y2y9 mus musculus
18	30	83.3	479	1 KLF_MOUSE	Q89091 mus musculus
19	30	83.3	480	1 KLF_MOUSE	Q13118 homo sapien
20	30	83.3	480	1 KLF_MOUSE	Q08876 rattus norv
21	30	83.3	511	1 KLF_MOUSE	Q9epf4 mus musculus
22	30	83.3	512	1 KLF_MOUSE	Q14901 homo sapien
23	30	83.3	606	1 SP2_HUMAN	Q02086 homo sapien
24	29	80.6	711	1 SP3_HUMAN	Q02447 homo sapien
25	29	80.6	203	1 YPC_CHL1	Q08263 chlorobium
26	29	80.6	298	1 MOX2_XENLA	P39021 xenopus lae
27	29	80.6	303	1 MOX2_HUMAN	P50222 homo sapien
28	29	80.6	303	1 MOX2_MOUSE	P32443 mus musculus
29	29	80.6	303	1 MOX2_RAT	P39020 rattus norv
30	29	80.6	496	1 PAL2_HUMAN	Q9upg8 homo sapien
31	29	80.6	864	1 FAS_YEAST	P53848 s folic aci
32	28	77.8	59	1 EGR1_SERCA	O73694 serinus can
33	28	77.8	62	1 EGR2_CRILLO	P26633 cricetus
34	28	77.8	62	1 EGR2_DUSTH	P26634 duscicyon th

RESULT 1

KLFF_MOUSE
ID KLFF_MOUSE STANDARD; PRT; 415: AA.
AC Q9EPW2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Krueppel-like factor 15 (Cardiovascular Kruppel-like factor).
GN KLF15 OR CKLF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RA Gray S.J., Kuo C.T., Leiden J.M., Jain M.K.;
RT "CKLF, a cardiovascular Kruppel-like factor.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
EX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -|- CELLULAR LOCATION: Nuclear.
CC -|- PROTEIN: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

34 28 77.8 62 1 EGR2_POERE
35 28 77.8 85 1 IYB2_BUNMU
36 28 77.8 193 1 EGR1_POEGU
37 28 77.8 194 1 EGR1_CHICK
38 28 77.8 194 1 EGR1_COTJA
39 28 77.8 199 1 KTHY_HALNI
40 28 77.8 204 1 GDIR_BOVIN
41 28 77.8 208 1 EDAD_MOUSE
42 28 77.8 219 1 EGL5_CAEEL
43 28 77.8 283 1 KLF6_HUMAN
44 28 77.8 283 1 KLF6_MOUSE
45 28 77.8 283 1 KLF6_RAT

ALIGNMENTS

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: AF317225; AAC38597.1; -
 CC EMBL: AK009739; BAB26470.1; -
 CC EMBL: BC013486; AAH13486.1; -
 CC HSSP: P08047; 1SP2.
 CC
 CC DR MGI:1929988; KLF15.
 CC InterPro: IPR000822; Znf_C2H2.
 CC Pfam: PF00096; zf-C2H2; 3.
 CC ProDom: PD000003; Znf_C2H2; 1.
 CC SMART: SM00355; Znf_C2H2; 3.
 CC ProSite: PS00028; ZINC_FINGER_C2H2_1; 3.
 CC ProSite: PS00157; ZINC_FINGER_C2H2_2; 3.
 CC Transcription regulation; DNA-binding; Activator; Nuclear protein;
 KW Repeat; Zinc-finger; Metal-binding.
 FT ZN_FING 320 344 C2H2-TYPE.
 FT ZN_FING 350 374 C2H2-TYPE.
 FT ZN_FING 380 402 C2H2-TYPE.
 SQ SEQUENCE 415 AA; 44252 MW; 127A7B80DB3E333CC CRC64;

Query Match 91.7%; Score 33; DB 1; Length 415;
 Best Local Similarity 85.7%; Pred. No. 5.3;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
 Db 391 RSDHLSK 397
 ID KLF_HUMAN STANDARD; PRT; 416 AA.

RESULT 2
 ID KLF_HUMAN STANDARD; PRT; 416 AA.
 AC Q5UIH9;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Kruppel-like factor 15 (Kidney-enriched kruppel-like factor).
 GN KLF15 OR KLF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
 RC TISSUE=Kidney;
 RX MEDLINE=20440192; PubMed=10982849;
 RA Uchida S., Tanaka Y., Ito H., Saitoh-Obara F., Inazawa J.,
 RA Yokoyama K.K., Sasaki S., Marumo F.;
 RT "Transcriptional regulation of the CLC-K1 promoter by myc-associated
 RT zinc finger protein and kidney-enriched Kruppel-like factor, a novel
 RT zinc finger repressor.";
 RL Mol. Cell. Biol. 20:7319-7331(2000).
 CC -1- FUNCTION: Transcriptional activator. Binds to the GA element of
 CC the CLCNKA promoter.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: Highly expressed in liver, followed by heart,
 CC skeletal muscle, and kidney. Not expressed in bone marrow or
 CC lymphoid tissues.
 CC -1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
 CC PROTEINS.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AB029254; BAA88561.1; -
 CC HSSP: P08047; 1SP2.

DR TRANSFAC; T05058; -
 DR Genew; HGNC:14536; KLF15.
 DR MIM; 606465; -
 DR InterPro: IPR000822; Znf_C2H2.
 DR Pfam: PF00096; zf-C2H2; 3.
 DR ProDom: PD000003; Znf_C2H2; 1.
 DR SMART: SM00355; Znf_C2H2; 3.
 DR ProSite: PS00028; ZINC_FINGER_C2H2_1; 3.
 DR ProSite: PS00157; ZINC_FINGER_C2H2_2; 3.
 DR Transcription regulation; DNA-binding; Activator; Nuclear protein;
 KW Repeat; Zinc-finger; Metal-binding.
 FT DOMAIN 321 403 ZINC_FINGERS.
 FT ZN_FING 321 345 C2H2-TYPE.
 FT ZN_FING 351 375 C2H2-TYPE.
 FT ZN_FING 381 403 C2H2-TYPE.
 SQ SEQUENCE 416 AA; 43992 MW; 6335F85141BEB276 CRC64;

Query Match 91.7%; Score 33; DB 1; Length 416;
 Best Local Similarity 85.7%; Pred. No. 5.3;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
 Db 392 RSDHLSK 398
 ID BT_DROME STANDARD; PRT; 644 AA.

RESULT 3
 ID BT_DROME STANDARD; PRT; 644 AA.
 AC Q24266; Q9W319;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transcription factor BT-D (Buttonhead protein).
 GN BT-D OR CG12653.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S; TISSUE=Embryo;
 RX MEDLINE=94081952; PubMed=8259212;
 RA Wimmer E.A., Jaecckle H., Pfeifle C., Cohen S.M.;
 RT "A Drosophila homologue of human Sp1 is a head-specific segmentation
 RT gene.";
 RL Nature 366:690-694(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Calniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Fouts C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "the genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: REQUIRED FOR THE DEVELOPMENT OF THE ANTENNA,
 CC INTERCALARY AND MANDIBULAR SEGMENTS OF THE HEAD.
 CC -1- SUBCELLULAR LOCATION: Nuclear (by similarity).
 CC -1- DEVELOPMENTAL STAGE: FIRST EXPRESSED IN A STRIPE COVERING THE HEAD
 CC ANLAGEN OF THE SYNCHITIAL BLASTODERM EMBRYO, PERSISTS THROUGH
 CC GASTRULATION AND DECAYS DURING GERM BAND EXTENSION. EXPRESSED
 CC LATER IN DEVELOPMENT IN A COMPLEX SPATIALLY RESTRICTED PATTERN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage -by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z29361; CAA82545.1; -;
 DR EMBL; AE003448; AAF46518.1; -;
 DR HSSP; P08047; 1SP2.
 DR FlyBase; FBgn0000233; btd.
 DR InterPro; IPR000822; Znf_C2H2.
 DR Pfam; PF000096; zf-C2H2; 3.
 DR PRINTS; PR00048; ZINCFINGER.
 DR ProDom; PD000003; Znf_C2H2; 1.
 DR SMART; SM00355; Znf_C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
 KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
 KW DNA-binding; Nuclear protein; Repeat.
 FT DOMAIN 333 413 ZINC FINGERS.
 FT ZN_FING 333 357 C2H2-TYPE.
 FT ZN_FING 363 385 C2H2-TYPE.
 FT ZN_FING 391 413 C2H2-TYPE.
 FT DOMAIN 14 89 GLN-RICH.
 FT DOMAIN 208 220 POLY-ALA.
 FT DOMAIN 431 434 POLY-ALA.
 FT DOMAIN 486 492 POLY-PRO.
 FT DOMAIN 499 502 POLY-THR.
 FT DOMAIN 515 519 POLY-SER.
 FT DOMAIN 530 536 POLY-SER.
 FT DOMAIN 596 599 POLY-SER.
 SQ SEQUENCE 644 AA; 68581 MW; A0DB98C2AF938452 CRC64;
 Query Match 91.7%; Score 33; DB 1; Length 644;
 Best Local Similarity 85.7%; Pred. No. 8.9;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RSDHLSR 7
 Db 402 RSDHLSK 408
 RESULT 4
 SP4_HUMAN
 ID SP4_HUMAN STANDARD; PRT; 784 AA.
 AC Q02446; O60402;

DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Transcription factor Sp4 (SPR-1).
 GN SP4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Uterus;
 RX MEDLINE=33087156; PubMed=1454515;
 RA Hagen G., Mueller S., Beato M., Suske G.;
 RT "Cloning by recognition site screening of two novel GT box binding
 RL proteins: a family of Sp1 related genes";
 RL Nucleic Acids Res. 20:5519-5525(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Ozersky P., Holmes A.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: BINDS TO GT AND GC BOXES PROMOTERS ELEMENTS. PROBABLE
 CC TRANSCRIPTIONAL ACTIVATOR.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: ABUNDANT IN BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
 CC PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage -by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X68561; CAA48563.1; -;
 DR EMBL; AC004595; AAD12226.1; -;
 DR PIR; S26638; S26638.
 DR HSSP; P08047; 1SP1.
 DR TRANSFAC; T02339; -;
 DR Gene; HGNC:11209; SP4.
 DR MIM; 600540; -;
 DR InterPro; IPR000822; Znf_C2H2.
 DR Pfam; PF000096; zf-C2H2; 3.
 DR PRINTS; PR00048; ZINCFINGER.
 DR ProDom; PD000003; Znf_C2H2; 2.
 DR SMART; SM00355; Znf_C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
 KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
 KW DNA-binding; Nuclear protein; Repeat.
 FT DOMAIN 7 11 POLY-GLU.
 FT DOMAIN 12 19 POLY-ALA.
 FT DOMAIN 122 130 POLY-SER.
 FT DOMAIN 185 188 POLY-SER.
 FT DOMAIN 647 729 ZINC FINGERS.
 FT ZN_FING 647 671 C2H2-TYPE.
 FT ZN_FING 677 701 C2H2-TYPE.
 FT ZN_FING 707 729 C2H2-TYPE.
 FT CONFLICT 197 197 K -> Q (IN REF. 2).
 FT CONFLICT 379 380 HA -> QP (IN REF. 2).
 FT CONFLICT 386 386 Q -> A (IN REF. 2).
 SQ SEQUENCE 784 AA; 82025 MW; 3C4EAE28CB28B1FB CRC64;
 Query Match 91.7%; Score 33; DB 1; Length 784;
 Best Local Similarity 85.7%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RSDHLSR 7
 Db 718 RSDHLSK 724

```

RESULT 5
SPL_HUMAN
ID SPL_HUMAN STANDARD; PRT; 785 AA.
AC P08047; Q9NVE7; Q9H3Q5;
DT 01-AUG-1988 (Rel. 08, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor Spl.
GN SPL OR TSFPL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 4-785 FROM N.A.
RC TISSUE=Cervical carcinoma;
RA Haggart M.H., Ladurner A.G.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-558 FROM N.A.
RX MEDLINE=20545561; PubMed=10973950;
RA Takahara T., Kanazu S., Yanagisawa S., Akanuma H.;
RT "heterogeneous Spl mRNAs in human HepG2 cells include a product of
RT homotypic trans-splicing.";
RL J. Biol. Chem. 275:38067-38072(2000).
RN [3]
RP SEQUENCE OF 90-785 FROM N.A., AND SEQUENCE OF 359-375 AND 670-675.
RX MEDLINE=88080466; PubMed=3319186;
RA Kadosaga J.T., Carner K.R., Masiarz F.R., Tjian R.;
RT "Isolation of cDNA encoding transcription factor Spl and functional
RT analysis of the DNA binding domain.";
RL Cell 51:1079-1090(1987).
RN [4]
RP O-GLYCOSYLATION.
RX MEDLINE=89003041; PubMed=3139301;
RA Jackson S.P., Tjian R.;
RT "O-glycosylation of eukaryotic transcription factors: implications
RT for mechanisms of transcriptional regulation.";
RL Cell 55:125-133(1988).
RN [5]
RP STRUCTURE BY NMR OF 654-684 AND 684-712.
RX MEDLINE=97218212; PubMed=9065444;
RA Narayan V.A., Krlwacki R.W., Caradonna J.P.;
RT "Structures of zinc finger domains from transcription factor Spl.
RT Insights into sequence-specific protein-DNA recognition.";
RL J. Biol. Chem. 272:7801-7809(1997).
RN [6]
RP IDENTIFICATION OF SEROTONIN 1A RECEPTOR PROMOTER BINDING SITES.
RX MEDLINE=96224025; PubMed=8626793;
RA Parks C.L., Shenk T.;
RT "The serotonin 1a receptor gene contains a TATA-less promoter that
RT responds to MAZ and Spl.";
RL J. Biol. Chem. 271:4417-4430(1996).
CC [1] FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
CC RECOGNITION SITES. CAN INTERACT WITH G/C-RICH MOTIFS FROM
CC SEROTONIN RECEPTOR PROMOTER.
CC [1] SUBCELLULAR LOCATION: Nuclear.
CC [1] PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS.
CC [1] SIMILARITY: BELONGS TO THE SPL FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF252284; AAF67726.1;
CC EMBL; AB039286; BAB13476.1;

```

```

DR EMBL; J03133; AAA61154.1;
DR PIR; A29635; A29635.
DR PDB; 1SP1; 21-APR-97.
DR PDB; 1SP2; 21-APR-97.
DR TRANSFAC; T00759;
DR GlycosuiteDB; P08047;
DR Genew; HGNC:11205; SPL.
DR MIM; 189906;
DR InterPro; IPR000822; znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR ProDom; PD000003; znf_C2H2; 2.
DR SMART; SM00355; znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
KW DNA-binding; Nuclear protein; Repeat; Glycoprotein; 3D-structure.
FT DOMAIN 626 708 ZINC FINGERS.
FT ZN_FING 626 650 C2H2-TYPE.
FT ZN_FING 656 680 C2H2-TYPE.
FT ZN_FING 686 708 C2H2-TYPE.
FT CONFLICT 366 366 D -> G (IN REF. 3; AA SEQUENCE).
FT CONFLICT 670 670 S -> F (IN REF. 3; AA SEQUENCE).
SQ SEQUENCE 785 AA; 80693 MW; 43893DBF6518B9EA CRC64;
Query Match 91.7%; Score 33; DB 1; Length 785;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 RSDHLRSR 7
DB 697 RSDHLRSK 703
RESULT 6
SPL_RAT
ID SPL_RAT STANDARD; PRT; 788 AA.
AC Q01714;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor Spl.
GN SPL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93010958; PubMed=1356762;
RA Imataka H., Sogawa K., Yasumoto K., Kikuchi Y., Sasano K.,
RA Kobayashi A., Hayami M., Fujii-Kuriyama Y.;
RT "Two regulatory proteins that bind to the basic transcription element
RT (BTE), a GC box sequence in the promoter region of the rat P-450IA1
RT gene.";
RL EMBO J. 11:3663-3671(1992).
CC [1] FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
CC RECOGNITION SITES.
CC [1] SUBCELLULAR LOCATION: Nuclear.
CC [1] PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS
CC (BY SIMILARITY).
CC [1] SIMILARITY: BELONGS TO THE SPL FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF252284; AAF67726.1;
CC EMBL; AB039286; BAB13476.1;

```


DR EMBL; D12768; BAA02235.1; -
DR PIR; JS0747; JS0747.
DR HSP; P08047; ISP1.
DR TRANSFAC; T00754; -
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 3.
KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
KW DNA-binding; Nuclear protein; Repeat; Glycoprotein.
FT DOMAIN 629 711 ZINC FINGERS.
FT ZN_FING 629 653 C2H2-TYPE.
FT ZN_FING 659 683 C2H2-TYPE.
FT ZN_FING 689 711 C2H2-TYPE.
SQ SEQUENCE 788 AA; 81015 MW; AA2B0CAB81AAB80C CRC64;

Query Match 91.7%; Score 33; DB 1; Length 788;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
|||||
Db 700 RSDHLSK 706

RESULT 7
KLFB_HUMAN
ID KLFB_HUMAN STANDARD; PRT; 359 AA.
AC O95600; Q9UGC4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Kruppel-like factor 8 (Zinc finger protein 741) (Basic kruppel-like factor 3).
GN KLFB OR ZNF741 OR BKLF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gorski J.L., MacDonald M., Vananthwerp M., Burright E.N., Bialecki M.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE-20221593; PubMed-10756197;
RA van Vliet J., Turner J., Crossley M.;
RT "Human Kruppel-like factor 8: a CACCC-box binding protein that associates with CtBP and represses transcription.";
RT Nucleic Acids Res. 28:1955-1962(2000).
RN [3]
RP SEQUENCE OF 1-299 FROM N.A.
RA Wray P.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Transcriptional repressor. Binds to CACCC-boxes promoter elements.
CC -!- SUBUNIT: Associates with CtBP.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: BELONGS TO THE SPI FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC

DR EMBL; U28282; AAC99849.1; -
DR EMBL; AL050309; CAB65785.1; -
DR HSP; P08047; ISP2.
DR Genew; HGNC:6351; KLF8.
DR MIM; 300286; -
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 3.
KW Transcription regulation; Repressor; Zinc-finger; Metal-binding;
KW DNA-binding; Nuclear protein; Repeat.
FT DOMAIN 274 356 ZINC FINGERS.
FT ZN_FING 274 298 C2H2-TYPE.
FT ZN_FING 304 328 C2H2-TYPE.
FT ZN_FING 334 356 C2H2-TYPE.
FT ZN_FING 356 363 E -> G (IN REF. 1).
SQ SEQUENCE 359 AA; 39313 MW; F8FDC1FD477C04F CRC64;

Query Match 86.1%; Score 31; DB 1; Length 359;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLS 6
|||||
Db 345 RSDHLS 350

RESULT 8
LCYB_SYNP7
ID LCYB_SYNP7 STANDARD; PRT; 411 AA.
AC Q55276;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lycopene beta cyclase (BC 1.14.-.-).
GN CRT1 OR LCY.
OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=1140;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95003701; PubMed-7919981;
RA Cunningham F.X. Jr., Sun Z., Chamovitz D., Hirschberg J., Gaunt E.;
RT "Molecular structure and enzymatic function of lycopene cyclase from the cyanobacterium Synecococcus sp strain PCC7942.";
RT Plant Cell 6:1107-1121(1994).
CC -!- FUNCTION: CATALYZES THE DOUBLE CYCLIZATION REACTION WHICH CONVERTS LYCOPENE TO BETA-CAROTENE AND NEUROSPORINE TO BETA-ZEACAROTENE.
CC -!- ENZYME REGULATION: INHIBITED BY THE BLEACHING HERBICIDE 2-(4-METHYLPHENOXY)TRIETHYLAMINE HYDROCHLORIDE (MPTA).
CC -!- PATHWAY: Carotenoid biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; X74599; CAA52677.1; -
DR InterPro; IPR000205; NAD_binding.
DR Oxidoreductase; NAD; Carotenoid biosynthesis.
FT NP_BIND 4 32 NAD (POTENTIAL).
SQ SEQUENCE 411 AA; 46085 MW; C46CC5B2E85E7AC2 CRC64;
Query Match 86.1%; Score 31; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLS 6

```
Db      181 RSDHLS 186
|||||
RESULT 9
SP6_MOUSE STANDARD; PRT; 152 AA.
AC Q9ESX2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transcription factor Sp6 (Krueppel-like factor 14) (Fragment).
GN SP6 OR KLF14.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20541714; PubMed=11087666;
RA Scohy S., Gabant P., Van Reeth T., Hertveldt V., Dreze P.-L.,
RA Van Vooren P., Riviere M., Szpirer J., Szpirer C.;
RT "Identification of KLF13 and KLF14 (SP6), novel members of the SP/XKLF
RT transcription factor family.";
RL Genomics 70:93-101(2000).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AJ275988; CAC06698.1; -.
DR HSSP; P08047; LSP2.
DR MGD; MGI:1932575; Sp6.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00056; zf-C2H2; 3.
DR ProDom; PD000003; Znf_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
DR Transcription regulation; DNA-binding; Nuclear protein; Repeat;
DR Zinc-finger; Metal-binding.
DR NON_TER
FT DOMAIN 30 112 ZINC FINGERS.
FT ZN_FING 30 54 C2H2-TYPE.
FT ZN_FING 60 84 C2H2-TYPE.
FT ZN_FING 90 112 C2H2-TYPE.
SQ SEQUENCE 152 AA; 16456 MW; AF629C4845599938 CRC64;

Query Match 83.3%; Score 30; DB 1; Length 152;
Best Local Similarity 71.4%; Pred. No. 7.7;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
|||||
Db 101 RSDHLAK 107

RESULT 10
BTBL_HUMAN STANDARD; PRT; 244 AA.
AC Q13886; Q16196;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transcription factor BTBL (Basic transcription element binding
```

DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transcription factor BTEB1 (Basic transcription element binding
DE protein 1) (BTE-binding protein 1) (GC box binding protein 1)
DE (Krueppel-like factor 9)
DE BTEB1 OR BTEB-1 OR BTEB OR KLF9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=99077958; PubMed=9858544;
RA Imhof A., Schuierer M., Werner O., Moser M., Roth C., Bauer R.,
RA Buettner R.;
RT "Transcriptional regulation of the AP-2alpha promoter by BTEB-1 and
RT AP-2rep, a novel wt-1/egr-related zinc finger repressor.";
RL Mol. Biol. 19:194-204(1999).
CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO GC BOX PROMOTER
CC ELEMENTS. SELECTIVELY ACTIVATES MRNA SYNTHESIS FROM GENES
CC CONTAINING TANDEM REPEATS OF GC BOXES BUT REPRESSES GENES WITH
CC A SINGLE GC BOX.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y14296; CAA74671.1; -
CC HSPF; P08047; ISP2.
CC MGD; MGI:1333856; Klf9.
CC InterPro; IPR000822; Znf_C2H2.
CC Pfam; PF00096; zf-C2H2; 3.
CC PRINTS; PR00048; ZINCfinger.
CC ProDom; PD000003; Znf_C2H2; 1.
CC SMART; SM00355; Znf_C2H2; 3.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW Transcription regulation; DNA-binding; Nuclear protein; Repeat;
KW Zinc-finger; Metal-binding.
FT DOMAIN 84 116 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 143 225 ZINC FINGERS.
FT ZN_FING 143 167 C2H2-TYPE.
FT ZN_FING 173 197 C2H2-TYPE.
FT ZN_FING 203 225 C2H2-TYPE.
SQ SEQUENCE 244 AA; 27169 MW; BDDF607FFA218D5A CRC64;
Query Match 83.3%; Score 30; DB 1; Length 244;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RSDHLRS 7
DB 214 RSDHLTK 220
RESULT 12
BTEB1_RAT
ID BTEB1_RAT STANDARD; PRT; 244 AA.
AC Q01713;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transcription factor BTEB1 (Basic transcription element binding
DE protein 1) (BTE-binding protein 1) (GC box binding protein 1).
GN BTEB1 OR BTEB.

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93010958; PubMed=1356762;
RA Imataka H., Sogawa K., Yasumoto K., Kikuchi Y., Sasano K.,
RA Kobayashi A., Hayami M., Fujii-Kuriyama Y.;
RT "Two regulatory proteins that bind to the basic transcription element
RT (BTE), a GC box sequence in the promoter region of the rat P-4501A1
RT gene.";
RL EMO J. 11:3663-3671(1992).
CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO GC BOX PROMOTER
CC ELEMENTS. SELECTIVELY ACTIVATES MRNA SYNTHESIS FROM GENES
CC CONTAINING TANDEM REPEATS OF GC BOXES BUT REPRESSES GENES WITH
CC A SINGLE GC BOX.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D12769; BAA02236.1; -
CC PIR; J50748; JS0748.
CC HSPF; P08047; ISP2.
CC TRANSFAC; T02210; -
CC InterPro; IPR000822; Znf_C2H2.
CC Pfam; PF00096; zf-C2H2; 3.
CC PRINTS; PR00048; ZINCfinger.
CC ProDom; PD000003; Znf_C2H2; 1.
CC SMART; SM00355; Znf_C2H2; 3.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW Transcription regulation; DNA-binding; Nuclear protein; Repeat;
KW Zinc-finger; Metal-binding.
FT DOMAIN 84 116 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 143 225 ZINC FINGERS.
FT ZN_FING 143 167 C2H2-TYPE.
FT ZN_FING 173 197 C2H2-TYPE.
FT ZN_FING 203 225 C2H2-TYPE.
SQ SEQUENCE 244 AA; 27155 MW; FBD1D13FEAFA37E0 CRC64;
Query Match 83.3%; Score 30; DB 1; Length 244;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RSDHLRS 7
DB 214 RSDHLTK 220
RESULT 13
BTE4_MOUSE
ID BTE4_MOUSE STANDARD; PRT; 251 AA.
AC P58334;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transcription factor BTEB4 (Basic transcription element binding-
DE protein 4) (BTE-binding protein 4) (Krueppel-like factor 16) (Dopamine
DE receptor regulating factor).
GN KLF16 OR BTEB4 OR DRFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=Neuroblastoma;
RX MEDLINE=21309923; PubMed=11390978;
RA Hwang C.K., D'Souza U.M., Elsch A.J., Yajima S., Lammers C.-H.,
RA Yang Y., Lee S.-H., Kim Y.-M., Nestler E.J., Mouradian M.M.;
RT "Dopamine receptor regulating factor, DRRF: a zinc finger
transcription factor.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7558-7563(2001).
CC -1- FUNCTION: Transcription factor that binds GC and GT boxes in the
CC d1A, D2 and D3 dopamine receptor promoters and displaces Sp1 and
CC Sp3 from these sequences. It modulates dopaminergic transmission
CC in the brain by repressing or activating transcription from
CC several different promoters depending on cellular context.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: High expression in brain; olfactory tubercle,
CC olfactory bulb, nucleus accumbens, striatum, hippocampal CA1
CC region, amygdala, dentate gyrus and frontal cortex. Moderate
CC expression in hippocampal CA2-3 regions, piriform cortex, septum,
CC and distinct thalamic nuclei. Low expression in the cerebellum.
CC -1- DOMAIN: The Ala/Pro-rich domain may contain discrete activation
CC and repression subdomains and also can mediate protein-protein
CC interactions.
CC -1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF283891; AAK66968.1; -
CC MGD: MG1:2153049; Klf16.
CC TRANSFAC: T05053; -
CC InterPro: IPR000822; Znf_C2H2.
CC Pfam: PF00096; zf-C2H2; 3.
CC SMART: SM00355; ZNF_C2H2; 3.
CC PROSITE: PS00028; ZINC_FINGER_C2H2_1; 3.
CC PROSITE: PS00157; ZINC_FINGER_C2H2_2; 3.
CC Transcription regulation; DNA-binding; Nuclear protein; Repeat;
CC Zinc-finger; Metal-binding.
CC FT DOMAIN 3 136 ALA/PRO-RICH.
FT DOMAIN 103 116 SER-RICH.
FT DOMAIN 126 208 ZINC FINGERS.
FT ZN_FING 126 150 C2H2-TYPE.
FT ZN_FING 156 180 C2H2-TYPE.
FT ZN_FING 186 208 C2H2-TYPE.
FT DOMAIN 223 248 PRO/SER-RICH.
SQ SEQUENCE 251 AA; 25665 MW; 3F0D7739B1A09FA4 CRC64;

Query Match 83.3%; Score 30; DB 1; Length 251;
Best Local Similarity 71.4%; Pred. No. 14;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLRSR 7
Db 197 RSDHLTK 203
|||||:
ID BTE4_HUMAN STANDARD; PRT; 252 AA.
AC Q9BXK1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Transcription factor BTE4 (Basic transcription element binding-
DE protein 4) (BTE-binding protein 4) (Krueppel-like factor 16) (Novel
DE Sp1-like zinc finger transcription factor 2) (Transcription factor
GN KLF16 OR BTE4 OR NSLP2).

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Conley A., Urrutia R.;
RT "Isolation of a novel zinc finger transcription factor from the
RT pancreas extends the repertoire of Spl-like proteins present in this
RT organ (Abstract #153).";
RL Pancreas 21:437-437(2000).
CC -1- FUNCTION: Transcription factor that binds GC and GT boxes and
CC displaces Sp1 and Sp3 from these sequences. Modulates dopaminergic
CC transmission in the brain (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF327440; AAK15698.1; -
CC HSSP: P08047; lSP2.
CC Genew: HGNC:16857; KLF16.
CC MIM: 606139; -
CC InterPro: IPR000822; Znf_C2H2.
CC Pfam: PF00096; zf-C2H2; 3.
CC SMART: SM00355; ZNF_C2H2; 3.
CC PROSITE: PS00028; ZINC_FINGER_C2H2_1; 3.
CC PROSITE: PS00157; ZINC_FINGER_C2H2_2; 3.
CC Transcription regulation; DNA-binding; Nuclear protein; Repeat;
CC Zinc-finger; Metal-binding.
CC FT DOMAIN 3 130 ALA/PRO-RICH.
FT DOMAIN 99 112 SER-RICH.
FT DOMAIN 127 209 ZINC FINGERS.
FT ZN_FING 127 150 C2H2-TYPE.
FT ZN_FING 157 181 C2H2-TYPE.
FT ZN_FING 187 209 C2H2-TYPE.
FT DOMAIN 210 249 PRO/SER-RICH.
SQ SEQUENCE 252 AA; 25430 MW; 9A0CB4B1A585A118 CRC64;

Query Match 83.3%; Score 30; DB 1; Length 252;
Best Local Similarity 71.4%; Pred. No. 14;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLRSR 7
Db 198 RSDHLAK 204
|||||:
ID KLF16_HUMAN STANDARD; PRT; 288 AA.
AC Q9Y2Y9; Q9Y356;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Krueppel-like factor 13 (Transcription factor BTEB3) (Basic
DE transcription element binding protein 3) (BTE-binding protein 3)
DE (KAT5 factor of late activated T lymphocytes-1) (KLFAT-1)
DE (Transcription factor NSLPI) (Novel Sp1-like zinc finger transcription
DE factor 1) (Transcription factor NSLP1).
GN KLF13 OR BTEB3 OR NSLP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

[1] SEQUENCE FROM N.A.
RX MEDLINE-95146379; PubMed=10023774;
RA Song A., Chen Y.F., Thamatrakoln K., Storm T.A., Krensky A.M.;
RT "RFLAT-1: a new zinc finger transcription factor that activates RANTES
gene expression in T lymphocytes.";
RL Immunity 10:93-103(1999).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE-Pancreas;
RX MEDLINE-99344343; PubMed=10415854;
RA Cook T., Gebelein B., Urrutia R.;
RT "Spl and its likes: biochemical and functional predictions for a
growing family of zinc finger transcription factors.";
RL Ann. N.Y. Acad. Sci. 880:94-102(1999).
[3]
RN SEQUENCE FROM N.A.
RC TISSUE-Lung;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
[4]
RN FUNCTION.
RX MEDLINE-21453359; PubMed=11477107;
RA Kaczynski J., Zhang J.S., Ellenrieder V., Conley A., Duenes T.,
RA Kester H., van Der Burg B., Urrutia R.;
RT "The Spl-like Protein BTEB3 inhibits transcription via the basic
transcription element box by interacting with msin3A and HDAC-1
co-repressors and competing with Spl.";
RL J. Biol. Chem. 276:36749-36756(2001).
CC -!- FUNCTION: Represses transcription by binding to the BTE site, a
GC-rich DNA element, in competition with the activator Spl. It
also represses transcription by interacting with the co-repressor
Sin3A and HDAC1. Activates RANTES expression in T cells.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- DOMAIN: The Ala/Pro-rich domain may contain discrete activation
and repression subdomains and also can mediate protein-protein
interactions.
CC -!- PTM: Phosphorylated.
CC -!- SIMILARITY: BELONGS TO THE SPL FAMILY OF C2H2-TYPE ZINC-FINGER
PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AF132599; AAD26864.1; -
DR EMBL; AF150628; AAD34020.1; -
DR EMBL; BC013946; AAH13946.1; -
DR EMBL; BC010438; AAH10438.1; -
DR EMBL; BC012741; AAH12741.1; -
DR HSSP; P08047; 1SP2.
DR TRANSFAC; T05051; -
DR Genew; HGNC:13672; KLF13.
DR MIM; 605328; -
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 6.
DR ProDom; PD000003; Znf_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW Transcription regulation; Activator; Repressor; DNA-binding;
KW Nuclear protein; Repeat; Zinc-finger; Metal-binding; Phosphorylation.
FT DOMAIN 2 145
FT DOMAIN 148 168 ARG/LYS-RICH (BASIC).
FT DOMAIN 167 249 ZINC FINGERS.
FT ZN_FING 167 191 C2H2-TYPE.
FT ZN_FING 197 221 C2H2-TYPE.
FT ZN_FING 227 249 C2H2-TYPE.

FT DOMAIN 264 287 SER-RICH.
FT DOMAIN 113 116 POLY-ALA.
FT CONFLICT 39 116 POLY-ALA.
FT CONFLICT 104 114 A -> S (IN REF. 2).
SQ SEQUENCE 288 AA; 31180 MW; DD2765EE00E9C049 CRC64;
Query Match 83.3%; Score 30; DB 1; Length 288;
Best Local Similarity 71.4%; Pred No. 16;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RSDHLSR 7
Db 238 RSDHLTK 244
|||||:
Search completed: July 15, 2003, 12:53:29
Job time : 7.66667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:50:51 ; Search time 12.3333 Seconds
(without alignments)
54.563 Million cell updates/sec

Title: US-10-006-069A-68

Perfect score: 36

Sequence: 1 RSDHLSR 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	1400	2 T22644	hypothetical prote
2	33	91.7	101	2 S30493	Spl protein - mous
3	33	91.7	644	2 S39356	transcription fact
4	33	91.7	696	2 A29635	transcription fact
5	33	91.7	780	2 A48143	HF-1 regulatory el
6	33	91.7	784	2 S26638	SPR-1 protein - hu
7	33	91.7	788	2 JS0747	regulatory protein
8	33	91.7	1464	2 JC5144	murinoglobulin pre
9	31	86.1	264	2 T27830	hypothetical prote
10	31	86.1	309	2 T25800	CH2-type zinc fin
11	31	86.1	497	2 T33634	hypothetical prote
12	30	83.3	233	2 T26781	hypothetical prote
13	30	83.3	244	2 I59602	GC box binding pro
14	30	83.3	244	2 JS0748	basic transcriptio
15	30	83.3	244	2 S25288	BTEB protein - rat
16	30	83.3	411	2 B83043	hypothetical prote
17	30	83.3	469	2 A57531	EGR alpha transcri
18	30	83.3	480	2 S23206	zinc finger protei
19	30	83.3	491	2 C83205	probable outer mem
20	30	83.3	495	2 A44489	GT box-binding pro
21	30	83.3	608	2 B70188	hypothetical prote
22	30	83.3	668	2 C24785	hypothetical prote
23	30	83.3	697	2 B44489	GT box-binding pro
24	30	83.3	1195	2 E96615	hypothetical prote
25	30	83.3	1355	2 T32092	hypothetical prote
26	30	83.3	1388	2 T30335	KLP2 protein - Afr
27	30	83.3	1400	2 T20904	hypothetical prote
28	30	83.3	1404	2 H89715	protein Fl4F4.3 [i
29	30	83.3	1427	2 T20903	hypothetical prote

30	29	80.6	103	2 T07578	hypothetical prote
31	29	80.6	165	2 T46963	sox4 protein [limp
32	29	80.6	298	2 S41469	homeotic protein M
33	29	80.6	302	2 A55641	homeotic protein G
34	29	80.6	303	2 A56837	homeotic protein M
35	29	80.6	303	2 B49122	homeobox protein M
36	29	80.6	303	2 A48130	growth arrest-spec
37	29	80.6	315	2 H87447	conserved hypothet
38	29	80.6	413	2 F96743	probable C2H2-type
39	29	80.6	609	2 A41081	alpha-1-inhibitor
40	29	80.6	864	1 S63229	probable multifunc
41	29	80.6	934	2 T47546	protein kinase-lik
42	29	80.6	2054	2 T32413	probable acetyl-Co
43	29	80.6	2251	2 T24490	hypothetical prote
44	28	77.8	59	1 TIKFBY	beta-1 bungarotoxi
45	28	77.8	61	1 TIKFB2	beta-2 bungarotoxi

ALIGNMENTS

RESULT 1

T22644

Hypothetical protein F54D1.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T22644

R:Lennard, N.

submitted to the EMBL Data Library, July 1996

A:Reference number: Z19592

A:Accession: T22644

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1400 <WIL>

A:Cross-references: EMBL:Z77132; PIDN:CAB00861.1; GSPDB:GN00022; CESP:F54D1.5

A:Experimental source: clone F54D1

C:Genetics:

A:Gene: CESP:F54D1.5

A:Map position: 4

A:Introns: 21/2; 51/2; 205/2; 276/3; 364/2; 466/3; 507/3; 536/3; 599/3; 672/2;

Query Match 100.0%; Score 36; DB 2; Length 1400;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSDHLSR 7

|||||||

Db 51 RSDHLSR 57

RESULT 2

S30493

Spl protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999

C:Accession: S30493

R:Chestier, A.; Charnay, P.

DNA Seq. 2, 325-327, 1992

A:Title: Difference in the genomic organizations of the related transcription factors

A:Reference number: S30493; MUID:92338398; PMID:1633330

A:Accession: S30493

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-101 <CH>

A:Cross-references: EMBL:X60136; NID:g54158; PIDN:CAA42721.1; PID:e38120; PID:gl33426

Query Match 91.7%; Score 33; DB 2; Length 101;

Best Local Similarity 85.7%; Pred. No. 2.5;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSDHLSR 7

|||||||

Db 15 RSDHLSR 21

```
HF-1 regulatory element binding protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A48143
R:Zhu, H.; Nguyen, V.T.; Brown, A.B.; Pourhossaini, A.; Garcia, A.V.; van Bilsen, M.;
MOL. Cell. Biol. 13, 4432-4444, 1993
A:Title: A novel, tissue-restricted zinc finger protein (HF-1b) binds to the cardiac
A:Reference number: A48143; MUID:93309478; PMID:8321243
A:Accession: A48143
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-780 <ZHU>
A:Experimental source: neonatal heart
A:Note: sequence extracted from NCBI backbone (NCBIN:134660, NCBIP:134661)

Query Match          91.7%; Score 33; DB 2; Length 780;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
    |||||
Db 714 RSDHLSK 720

RESULT 6
SPR-1 protein - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 08-Oct-1999
C:Accession: S26638
R:Hagen, G.; Mueller, S.; Beato, M.; Suske, G.
Nucleic Acids Res. 20, 5519-5525, 1992
A:Title: Cloning by recognition site screening of two novel GT box binding proteins:
A:Reference number: S26638; MUID:93087156; PMID:1454515
A:Accession: S26638
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-784 <HAG>
A:Cross-references: EMBL:X68561; NID:g38419; PIDN:CAA48563.1; PID:g38420
C:Genetics:
A:Gene: GDB:SP4: SPR-1
A:Cross-references: GDB:136781
A:Map position: 2q31-2q31
C:Keywords: DNA binding; transcription regulation

Query Match          91.7%; Score 33; DB 2; Length 784;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
    |||||
Db 718 RSDHLSK 724

RESULT 7
JS0747
regulatory protein Sp1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 01-Dec-2000
C:Accession: JS0747; S25287
R:Imataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Haya
submitted to JIPID, September 1992
A:Reference number: JS0747
A:Accession: JS0747
A:Molecule type: mRNA
A:Residues: 1-788 <IMA>
A:Cross-references: DDBJ:D12768; NID:g220911; PIDN:BAA02235.1; PID:d1002730; PID:g220
R:Imataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Haya
EMBO J. 11, 3663-3671, 1992
A:Title: Two regulatory proteins that bind to the basic transcription element (BTE),
A:Reference number: S25287; MUID:93010958; PMID:1356762
A:Accession: S25287
A:Molecule type: mRNA
```


A:Residues: 1-122,'L',124-311,'A',313-788 <IM2>
C:Keywords: DNA binding; transcription regulation

Query Match 91.7%; Score 33; DB 2; Length 788;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
|||||
Db 700 RSDHLSR 706

RESULT 8

JC5144
murinoglobulin precursor - guinea pig
C:Species: Cavia porcellus (guinea pig)
C:Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 17-Mar-1999
C:Accession: JC5144
R:Iwasaki, H.; Suzuki, Y.; Sinohara, H.
J. Biochem. 120, 1167-1175, 1996
A:Title: Cloning and sequencing of cDNAs encoding plasma alpha-macroglobulin and murinog
A:Reference number: JC5143; MUID:97164019; PMID:9010766
A:Contents: liver
A:Accession: JC5144
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1464 <IWA>
A:Cross-references: DDBJ:D84339
C:Superfamily: alpha-2-macroglobulin
F:1-23/Domain: signal sequence #status predicted <STG>

Query Match 91.7%; Score 33; DB 2; Length 1464;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
|||||
Db 1392 RSDHVS 1398

RESULT 9

T27830
hypothetical protein ZK337.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27830
R:White, S.

submitted to the EMBL Data Library, November 1996
A:Reference number: Z20426
A:Accession: T27830
A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA
A:Residues: 1-264 <WIL>
A:Cross-references: EMBL:Z82090; PIDN:CAB05008.1; GSPDB:GN00019; CESP:ZK337.2
A:Experimental source: clone ZK337
C:Genetics:
A:Gene: CESP:ZK337.2
A:Map position: 1
A:Introns: 60/3; 107/2; 155/3

Query Match 86.1%; Score 31; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLS 6
|||||
Db 129 RSDHLS 134

RESULT 10

T25800
C2H2-type zinc finger domain, Wt1 homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25800
R:Fulton, L.

submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid F54H5.
A:Reference number: Z20089

A:Accession: T25800
A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA
A:Residues: 1-309 <FUL>

A:Cross-references: EMBL:U80952; PIDN:AAB38094.1; GSPDB:GN00020; CESP:mua-1
A:Experimental source: strain Bristol N2; clone F54H5
C:Genetics:
A:Gene: CESP:mua-1
A:Map position: 2
A:Introns: 31/1; 178/2; 234/2; 273/2

Query Match 86.1%; Score 31; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLS 6
|||||
Db 295 RSDHLS 300

RESULT 11

T33634
hypothetical protein F56F11.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33634

R:Latreille, P.; Kramer, J.; Keppeler, D.
submitted to the EMBL Data Library, October 1998

A:Description: The sequence of C. elegans cosmid F56F11.
A:Reference number: Z21379
A:Accession: T33634

A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-497 <LAT>

A:Cross-references: EMBL:AF099922; PIDN:AAC71780.1; GSPDB:GN00021; CESP:F56F11.3
A:Experimental source: strain Bristol N2; clone F56F11
C:Genetics:
A:Gene: CESP:F56F11.3
A:Map position: 3
A:Introns: 46/1; 103/3; 168/1; 195/1; 294/1; 440/1; 470/1

Query Match 86.1%; Score 31; DB 2; Length 497;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLS 6
|||||
Db 485 RSDHLS 490

RESULT 12

T26781

hypothetical protein Y40B1A.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26781

R:Harris, B.
submitted to the EMBL Data Library, December 1998

A:Reference number: Z20264
A:Accession: T26781
A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA
A:Residues: 1-233 <WIL>

A:Cross-references: EMBL:AL034392; PIDN:CRA22307.1; CESP:Y40B1A.4
A:Experimental source: clone Y40B1A
C:Genetics:
A:Gene: CESP:Y40B1A.4

A; Introns: 13/1; 45/3; 99/2; 178/2

Query Match 83.3%; Score 30; DB 2; Length 233;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 206 RSDHLTK 212
|||||

RESULT 13

. I59602
GC box binding protein - human
C; Species: Homo sapiens (man)
C; Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 01-Dec-2000
C; Accession: I59602; I55409
R; Ohe, N.; Yamasaki, Y.; Sogawa, K.; Inazawa, J.; Ariyama, T.; Oshimura, M.; Fujii-Kuriyama, Y.; Sasano, K.; Kobayashi, A.; Haya-
Sonomat. Cell Mol. Genet. 19, 499-503, 1993
A; Title: Chromosomal localization and cDNA sequence of human BTEB, a GC box binding protein
A; Reference number: I59602; MUID: 94120483; PMID: 8291025
A; Accession: I59602
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: mRNA
A; Residues: 1-244 <RES>
A; Cross-references: GB:D31716; NID: g505081; PIDN: BAA06524.1; PID: g1060891
R; Imataka, H.; Nakayama, K.; Yasumoto, K.; Mizuno, A.; Fujii-Kuriyama, Y.; Hayami, M.
J. Biol. Chem. 269, 20668-20673, 1994
A; Title: Cell-specific transcriptional control of transcription factor BTEB expression. The
A; Reference number: I55409; MUID: 94327649; PMID: 8051167
A; Accession: I55409
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: mRNA
A; Residues: 1-31 <RE2>
A; Cross-references: GB: S72504; NID: g619341; PIDN: RAD14110.1; PID: g4261810
C; Genetics:
A; Gene: BTEB

Query Match 83.3%; Score 30; DB 2; Length 244;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 214 RSDHLTK 220
|||||

RESULT 14

JS0748
basic transcription element-binding protein - rat
C; Species: Rattus norvegicus (Norway rat)
C; Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 05-Nov-1999
C; Accession: JS0748
R; Imataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Hayami,
submitted to JIPID, September 1992
A; Reference number: JS0747
A; Accession: JS0748
A; Molecule type: mRNA
A; Residues: 1-244 <IMA>
A; Cross-references: DDBJ: D12769; NID: g220677; PIDN: BAA02236.1; PID: d1002731; PID: g220678
C; Keywords: DNA binding; transcription regulation

Query Match 83.3%; Score 30; DB 2; Length 244;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 214 RSDHLTK 220
|||||

RESULT 15

S25288

BTEB-protein - rat

C; Species: Rattus norvegicus (Norway rat)
C; Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
R; Imataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Haya-
EMBO J. 11, 3663-3671, 1992
A; Title: Two regulatory proteins that bind to the basic transcription element (BTE),
A; Reference number: S25287; MUID: 93010958; PMID: 1356762
A; Accession: S25288
A; Molecule type: mRNA
A; Residues: 1-244 <IMA>

Query Match 83.3%; Score 30; DB 2; Length 244;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 214 RSDHLTK 220
|||||

Search completed: July 15, 2003, 12:55:30
Job time: 13.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:48:06 : Search time 33.3333 Seconds
(without alignments)
27.983 Million cell updates/sec

Title: US-10-006-069a-55
Perfect score: 35
Sequence: 1 DRSNLTR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	7	22 AAB84241	Zinc protein recog
2	35	100.0	7	23 AAB83564	F2 zinc finger for
3	35	100.0	7	23 AAB83793	Human VEGF-targete
4	35	100.0	7	23 AAB83812	Human VEGF-targete
5	35	100.0	7	23 AAB83813	Human VEGF-targete
6	35	100.0	7	23 AAB83816	Human VEGF-targete
7	35	100.0	7	23 AAB83824	Human VEGF-targete
8	35	100.0	7	23 AAB83888	Human VEGF-targete
9	35	100.0	7	23 AAB83910	Human VEGF-targete
10	35	100.0	7	23 AAB83913	Human VEGF-targete

11	35	100.0	7	23 AAB84241	Rat VEGF-targeted
12	35	100.0	7	23 AAB84241	Rat VEGF-targeted
13	35	100.0	7	23 AAB84241	Human ER-alpha loc
14	35	100.0	7	23 AAB84241	Human ER-alpha loc
15	35	100.0	7	23 AAB84241	Zinc finger protei
16	35	100.0	7	23 AAB84241	Zinc finger protei
17	35	100.0	7	23 AAB84241	Zinc finger protei
18	35	100.0	7	23 AAB84241	Zinc finger protei
19	35	100.0	7	23 AAB84241	Zinc finger protei
20	35	100.0	7	23 AAB84241	Zinc finger protei
21	35	100.0	7	23 AAB84241	Zinc finger protei
22	35	100.0	7	23 AAB84241	Zinc finger protei
23	35	100.0	7	23 AAB84241	Zinc finger protei
24	35	100.0	7	23 AAB84241	Zinc finger protei
25	35	100.0	7	23 AAB84241	Zinc finger protei
26	35	100.0	7	23 AAB84241	Zinc finger protei
27	35	100.0	7	23 AAB84241	Zinc finger protei
28	35	100.0	7	23 AAB84241	Zinc finger protei
29	35	100.0	7	23 AAB84241	Zinc finger protei
30	35	100.0	7	23 AAB84241	Zinc finger protei
31	35	100.0	7	23 AAB84241	Zinc finger protei
32	35	100.0	7	23 AAB84241	Zinc finger protei
33	35	100.0	7	23 AAB84241	Zinc finger protei
34	35	100.0	7	23 AAB84241	Zinc finger protei
35	35	100.0	7	23 AAB84241	Zinc finger protei
36	35	100.0	7	23 AAB84241	Zinc finger protei
37	35	100.0	7	23 AAB84241	Zinc finger protei
38	35	100.0	7	23 AAB84241	Zinc finger protei
39	35	100.0	7	23 AAB84241	Zinc finger protei
40	35	100.0	7	23 AAB84241	Zinc finger protei
41	35	100.0	7	23 AAB84241	Zinc finger protei
42	35	100.0	7	23 AAB84241	Zinc finger protei
43	35	100.0	7	23 AAB84241	Zinc finger protei
44	35	100.0	7	23 AAB84241	Zinc finger protei
45	35	100.0	7	23 AAB84241	Zinc finger protei

ALIGNMENTS

RESULT 1
AAB84241
ID AAB84241 standard; peptide; 7 AA.
XX
AC AAB84241;
XX
DT 06-AUG-2001 (first entry)
XX
DE Zinc protein recognition helix SBS9 for target DNA triplet GAC.
XX
KW Phenotype associated gene; zinc finger protein; cancer; nephritis;
KW prostate hypertrophy; hematopoiesis; osteoporosis; obesity;
KW cardiovascular disease; diabetes.
XX
XX Synthetic.
XX OS
XX WO200140798-A2.
XX PN
XX 07-JUN-2001.
XX PD
XX 06-DEC-2000; 2000WO-US33086.
XX PF
XX 06-DEC-1999; 99US-0456100.
XX PR
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX FA
XX PI
XX Case CC, Liu Q, Rebar EJ;
XX WPI; 2001-374953/39.
XX
XX Identifying genes associated with selected phenotype for research
XX purposes, involves culturing cells transduced with nucleic acid
XX encoding zinc finger proteins and assaying cells exhibiting selected

PF phenotype -
 XX Example 1; Page 36; 58pp; English.
 XX
 CC The specification describes a method for identifying genes associated
 CC with a selected phenotype. The method involves providing a library of
 CC nucleotide sequences encoding partially randomized zinc finger proteins,
 CC transducing cells with expression vectors, each comprising a sequence
 CC from the library, culturing the cells for expressing the zinc finger
 CC protein, assaying the cells for selected phenotype, and identifying the
 CC gene of interest, in cells exhibiting the phenotype. The method is useful
 CC for identifying a gene or genes associated with a selected phenotype such
 CC as the one related to cancer, nephritis, prostate hypertrophy,
 CC hematopoiesis, osteoporosis, obesity, cardiovascular disease or diabetes.
 CC The method is useful in academic laboratories, in the biotechnological
 CC industries, and in pharmaceutical, genomic, agricultural and chemical
 CC companies. AAB84233-44 represent recognition helices of zinc finger
 CC proteins, which recognise different DNA triplets.
 XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 35; DB 22; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRSNLTNR 7
 Db 1 DRSNLTNR 7
 |||||
 RESULT 2
 ABB83564
 ID ABB83564 standard; peptide; 7 AA.
 XX
 AC ABB83564;
 XX
 DT 27-SEP-2002 (first entry)
 XX
 DE F2 zinc finger for target sequence ZFP 1.
 XX
 KW zinc finger; stress tolerance; pathogen resistance;
 KW agrochemical.
 XX
 OS Unidentified.
 XX
 PN WO200257294-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 22-JAN-2002; 2002WO-US01906.
 XX
 PR 22-JAN-2001; 2001US-263445P.
 PR 11-MAY-2001; 2001US-290716P.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 PI Jamieson A, Li G;
 XX
 DR WPI; 2002-566792/60.
 XX
 PT Modified plant zinc finger protein for modulating gene expression in a
 PT plant cell comprises zinc fingers that bind to a target site -
 XX
 XX Example 4; Page 42; 50pp; English.
 XX
 CC The present invention relates to a modified plant zinc finger
 CC protein. This zinc finger protein is used to modulated gene
 CC expression in a plant cell. Nucleic acid encoding the zinc finger is
 CC expressed in plant cells to produce a plant with an altered phenotype
 CC relative to the wild-type plant. The altered phenotype is high in
 CC nutritional value, yield, stress tolerance, pathogen resistance,
 CC resistance to agrochemicals, production of pharmaceutical compounds or
 CC reduction of industrial chemicals. The present sequence is

CC a zinc finger protein sequence that is attracted to a ZFP target
 CC sequence.
 XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 35; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRSNLTNR 7
 Db 1 DRSNLTNR 7
 |||||
 RESULT 3
 ABB03793
 ID ABB03793 standard; Peptide; 7 AA.
 XX
 AC ABB03793;
 XX
 DT 25-SEP-2002 (first entry)
 XX
 DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 36.
 XX
 KW zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
 KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
 KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnery;
 KW antitumor; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
 KW osteopathic; antiinfertility.
 XX
 OS Homo sapiens.
 XX
 PN WO200246412-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 06-DEC-2001; 2001WO-US46861.
 XX
 PR 07-DEC-2000; 2000US-0733604.
 PR 12-DEC-2000; 2000US-0736083.
 PR 30-APR-2001; 2001US-0846033.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
 PI Jarvis E;
 XX
 DR WPI; 2002-527918/56.
 XX
 PT New zinc finger protein that binds to target site in vascular
 PT endothelial growth factor gene, useful for modulating expression of the
 PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
 PT ulcer -
 XX
 PS Claim 4; Page 102; 195pp; English.
 XX
 CC The present invention relates to a zinc finger protein that binds to a
 CC target site in one or more vascular endothelial growth factor (VEGF)
 CC genes. The protein is useful for modulating expression of a VEGF gene,
 CC thereby regulating angiogenesis and vasculogenesis. This can be used to
 CC treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
 CC diabetic retinopathy or psoriasis. The present sequence is a peptide
 CC shown in the invention.
 XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 35; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRSNLTNR 7
 Db 1 DRSNLTNR 7
 |||||

RESULT 4

ABJ03812
ID ABJ03812 standard; Peptide; 7 AA.

XX AC ABJ03812;
XX AC
DT 25-SEP-2002 (first entry)
XX Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 55.
DE
KW zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnary;
KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
KW osteopathic; antiinfertility.

XX Homo sapiens.

XX WO200246412-A2.

XX 13-JUN-2002.

XX 06-DEC-2001; 2001WO-US46861.

XX 07-DEC-2000; 2000US-0733604.

XX 12-DEC-2000; 2000US-0736083.

XX 30-APR-2001; 2001US-0846033.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
XX Jarvis E;

XX WPI; 2002-527918/56.

XX New zinc finger protein that binds to target site in vascular
XX endothelial growth factor gene, useful for modulating expression of the
XX gene and for treating atherosclerosis, ischemia, arthritis, wound or
XX ulcer

XX Claim 4; Page 103; 195pp; English.

XX The present invention relates to a zinc finger protein that binds to a
XX target site in one or more vascular endothelial growth factor (VEGF)
XX genes. The protein is useful for modulating expression of a VEGF gene,
XX thereby regulating angiogenesis and vasculogenesis. This can be used to
XX treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
XX diabetic retinopathy or psoriasis. The present sequence is a peptide
XX shown in the invention.

XX Sequence 7 AA;

Query Match 100.0%; Score 35; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7

DB 1 DRSNLTR 7

RESULT 5

ABJ03813
ID ABJ03813 standard; Peptide; 7 AA.

XX AC ABJ03813;

XX 25-SEP-2002 (first entry)

XX Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 56.

XX

KW zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnary;
KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
KW osteopathic; antiinfertility.

XX Homo sapiens.

XX WO200246412-A2.

XX 13-JUN-2002.

XX 06-DEC-2001; 2001WO-US46861.

XX 07-DEC-2000; 2000US-0733604.

XX 12-DEC-2000; 2000US-0736083.

XX 30-APR-2001; 2001US-0846033.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
XX Jarvis E;

XX WPI; 2002-527918/56.

XX New zinc finger protein that binds to target site in vascular
XX endothelial growth factor gene, useful for modulating expression of the
XX gene and for treating atherosclerosis, ischemia, arthritis, wound or
XX ulcer

XX Claim 4; Page 103; 195pp; English.

XX The present invention relates to a zinc finger protein that binds to a
XX target site in one or more vascular endothelial growth factor (VEGF)
XX genes. The protein is useful for modulating expression of a VEGF gene,
XX thereby regulating angiogenesis and vasculogenesis. This can be used to
XX treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
XX diabetic retinopathy or psoriasis. The present sequence is a peptide
XX shown in the invention.

XX Sequence 7 AA;

Query Match 100.0%; Score 35; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7

DB 1 DRSNLTR 7

RESULT 6

ABJ03816
ID ABJ03816 standard; Peptide; 7 AA.

XX AC ABJ03816;

XX 25-SEP-2002 (first entry)

XX Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 59.

XX

KW zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnary;
KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
KW osteopathic; antiinfertility.

XX Homo sapiens.

XX WO200246412-A2.

XX 13-JUN-2002.

XX

PF 06-DEC-2001; 2001WO-US46861.
 XX
 PR 07-DEC-2000; 2000US-0733604.
 PR 12-DEC-2000; 2000US-0736083.
 PR 30-APR-2001; 2001US-0846033.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
 PI Jarvis E;
 XX WPI; 2002-527918/56.
 DR
 XX New zinc finger protein that binds to target site in vascular
 PT endothelial growth factor gene, useful for modulating expression of the
 PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
 PT ulcer -
 XX
 PS Claim 4; Page 102; 195pp; English.
 XX
 CC The present invention relates to a zinc finger protein that binds to a
 CC target site in one or more vascular endothelial growth factor (VEGF)
 CC genes. The protein is useful for modulating expression of a VEGF gene,
 CC thereby regulating angiogenesis and vasculogenesis. This can be used to
 CC treat atherosclerosis, ischemia, arthritis, wounds, ulcers, tumours,
 CC diabetic retinopathy or psoriasis. The present sequence is a peptide
 CC shown in the invention.
 XX
 SQ Sequence 7 AA;
 XX
 Query Match 100.0%; Score 35; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRSNLTR 7
 Db 1 DRSNLTR 7
 RESULT 7
 ABJ03824
 ID ABJ03824 standard; Peptide; 7 AA.
 XX
 AC ABJ03824;
 XX
 DT 25-SEP-2002 (first entry)
 XX
 DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 67.
 XX
 KW Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
 KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
 KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnerary;
 KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
 KW osteopathic; antinfertility.
 XX
 OS Homo sapiens.
 XX
 PN WO200246412-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 06-DEC-2001; 2001WO-US46861.
 XX
 PR 07-DEC-2000; 2000US-0733604.
 PR 12-DEC-2000; 2000US-0736083.
 PR 30-APR-2001; 2001US-0846033.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
 PI Jarvis E;
 XX WPI; 2002-527918/56.
 DR

XX New zinc finger protein that binds to target site in vascular
 PT endothelial growth factor gene, useful for modulating expression of the
 PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
 PT ulcer -
 XX
 PS Claim 4; Page 102; 195pp; English.
 XX
 CC The present invention relates to a zinc finger protein that binds to a
 CC target site in one or more vascular endothelial growth factor (VEGF)
 CC genes. The protein is useful for modulating expression of a VEGF gene,
 CC thereby regulating angiogenesis and vasculogenesis. This can be used to
 CC treat atherosclerosis, ischemia, arthritis, wounds, ulcers, tumours,
 CC diabetic retinopathy or psoriasis. The present sequence is a peptide
 CC shown in the invention.
 XX
 SQ Sequence 7 AA;
 XX
 Query Match 100.0%; Score 35; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRSNLTR 7
 Db 1 DRSNLTR 7
 RESULT 8
 ABJ03888
 ID ABJ03888 standard; Peptide; 7 AA.
 XX
 AC ABJ03888;
 XX
 DT 25-SEP-2002 (first entry)
 XX
 DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 134.
 XX
 KW Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
 KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
 KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnerary;
 KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
 KW osteopathic; antinfertility.
 XX
 OS Homo sapiens.
 XX
 PN WO200246412-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 06-DEC-2001; 2001WO-US46861.
 XX
 PR 07-DEC-2000; 2000US-0733604.
 PR 12-DEC-2000; 2000US-0736083.
 PR 30-APR-2001; 2001US-0846033.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
 PI Jarvis E;
 XX WPI; 2002-527918/56.
 DR
 XX New zinc finger protein that binds to target site in vascular
 PT endothelial growth factor gene, useful for modulating expression of the
 PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
 PT ulcer -
 XX
 PS Claim 6; Page 103; 195pp; English.
 XX
 CC The present invention relates to a zinc finger protein that binds to a
 CC target site in one or more vascular endothelial growth factor (VEGF)
 CC genes. The protein is useful for modulating expression of a VEGF gene,
 CC thereby regulating angiogenesis and vasculogenesis. This can be used to

CC treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
 CC diabetic retinopathy or psoriasis. The present sequence is a peptide
 CC shown in the invention.

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
 DB 1 DRSNLTR 7
 |||||

RESULT 9
 ABJ03910
 ID ABJ03910 standard; Peptide; 7 AA.
 XX AC ABJ03910;
 XX XX
 XX 25-SEP-2002 (first entry)
 XX XX
 DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 167.
 XX XX
 KW zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
 KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
 KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnerary;
 KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
 KW osteopathic; antiinfertility.
 XX OS Homo sapiens.
 XX XX
 PN WO200246412-A2.
 XX XX
 PD 13-JUN-2002.
 XX XX
 PF 06-DEC-2001; 2001WO-US46861.
 XX XX
 PR 07-DEC-2000; 2000US-0733604.
 PR 12-DEC-2000; 2000US-0736083.
 PR 30-APR-2001; 2001US-0846033.
 XX XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX XX
 PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
 PI Jarvis E;
 XX XX
 DR WPI; 2002-527918/56.
 XX XX
 PF New zinc finger protein that binds to target site in vascular
 PT endothelial growth factor gene, useful for modulating expression of the
 PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
 PT ulcer -
 XX XX
 PS Example 1; Page 104; 195pp; English.
 XX XX
 CC The present invention relates to a zinc finger protein that binds to a
 CC target site in one or more vascular endothelial growth factor (VEGF)
 CC genes. The protein is useful for modulating expression of a VEGF gene,
 CC thereby regulating angiogenesis and vasculogenesis. This can be used to
 CC treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
 CC diabetic retinopathy or psoriasis. The present sequence is a peptide
 CC shown in the invention.

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
 DB 1 DRSNLTR 7
 |||||

RESULT 11
 ABJ03935
 ID ABJ03935 standard; Peptide; 7 AA.
 XX AC ABJ03935;
 XX XX
 XX 25-SEP-2002 (first entry)
 XX XX
 DE Rat VEGF-targeted zinc finger protein fragment SEQ ID NO: 196.

DB 1 DRSNLTR 7

RESULT 10
 ABJ03913
 ID ABJ03913 standard; Peptide; 7 AA.
 XX AC ABJ03913;
 XX XX
 XX 25-SEP-2002 (first entry)
 XX XX
 DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 170.
 XX XX
 KW zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
 KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
 KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnerary;
 KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
 KW osteopathic; antiinfertility.
 XX OS Homo sapiens.
 XX XX
 PN WO200246412-A2.
 XX XX
 PD 13-JUN-2002.
 XX XX
 PF 06-DEC-2001; 2001WO-US46861.
 XX XX
 PR 07-DEC-2000; 2000US-0733604.
 PR 12-DEC-2000; 2000US-0736083.
 PR 30-APR-2001; 2001US-0846033.
 XX XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX XX
 PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
 PI Jarvis E;
 XX XX
 DR WPI; 2002-527918/56.
 XX XX
 PF New zinc finger protein that binds to target site in vascular
 PT endothelial growth factor gene, useful for modulating expression of the
 PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
 PT ulcer -
 XX XX
 PS Example 1; Page 104; 195pp; English.
 XX XX
 CC The present invention relates to a zinc finger protein that binds to a
 CC target site in one or more vascular endothelial growth factor (VEGF)
 CC genes. The protein is useful for modulating expression of a VEGF gene,
 CC thereby regulating angiogenesis and vasculogenesis. This can be used to
 CC treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
 CC diabetic retinopathy or psoriasis. The present sequence is a peptide
 CC shown in the invention.

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
 DB 1 DRSNLTR 7
 |||||

RESULT 11
 ABJ03935
 ID ABJ03935 standard; Peptide; 7 AA.
 XX AC ABJ03935;
 XX XX
 XX 25-SEP-2002 (first entry)
 XX XX
 DE Rat VEGF-targeted zinc finger protein fragment SEQ ID NO: 196.

XX Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
 KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
 KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnary;
 KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
 KW osteopathic; antiinfertility.
 XX Rattus sp.
 XX WO200246412-A2.
 XX 13-JUN-2002.
 XX 06-DEC-2001; 2001WO-US46861.
 XX 07-DEC-2000; 2000US-0733604.
 PR 12-DEC-2000; 2000US-0736083.
 PR 30-APR-2001; 2001US-0846033.
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 XX Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
 PI Jarvis E;
 PI WPI; 2002-527918/56.
 XX New zinc finger protein that binds to target site in vascular
 PT endothelial growth factor gene, useful for modulating expression of the
 PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
 PT ulcer -
 XX Disclosure; Page 105; 195pp; English.
 XX The present invention relates to a zinc finger protein that binds to a
 CC target site in one or more vascular endothelial growth factor (VEGF)
 CC genes. The protein is useful for modulating expression of a VEGF gene,
 CC thereby regulating angiogenesis and vasculogenesis. This can be used to
 CC treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
 CC diabetic retinopathy or psoriasis. The present sequence is a peptide
 CC shown in the invention.
 XX Sequence 7 AA;
 SQ Query Match 100.0%; Score 35; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRSNLTR 7
 Db |||||||
 1 DRSNLTR 7
 RESULT 12
 ABJ03937
 ID ABJ03937 standard; Peptide; 7 AA.
 AC ABJ03937;
 XX 25-SEP-2002 (first entry)
 DT Human ER-alpha locus targeting ZFP1 peptide #8.
 DE ZFP; cytostatic; antidiabetic; ophthalmological; vasotropic; chromatin;
 KW gene expression; antirheumatic; antiarthritic; antipsoriatic; nootropic;
 KW neuroprotective; cerebroprotective; estrogen receptor alpha; ER-alpha;
 KW zinc finger protein.
 XX Synthetic.
 OS WO200244386-A2.
 PN 06-JUN-2002.
 XX 30-NOV-2001; 2001WO-US45098.
 PF 01-DEC-2000; 2000US-250804P.
 PR (SANG-) SANGAMO BIOSCIENCES INC.
 PA Wolffe AP, Tse C, Collingwood T;
 XX WPI; 2002-537455/57.
 DR Regulating expression of gene by contacting cell with regulatory
 PT molecule comprising DNA-binding domain targeted to sequence within

XX 06-DEC-2001; 2001WO-US46861.
 XX 07-DEC-2000; 2000US-0733604.
 PR 12-DEC-2000; 2000US-0736083.
 PR 30-APR-2001; 2001US-0846033.
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 XX Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
 PI Jarvis E;
 PI WPI; 2002-527918/56.
 XX New zinc finger protein that binds to target site in vascular
 PT endothelial growth factor gene, useful for modulating expression of the
 PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
 PT ulcer -
 XX Disclosure; Page 105; 195pp; English.
 XX The present invention relates to a zinc finger protein that binds to a
 CC target site in one or more vascular endothelial growth factor (VEGF)
 CC genes. The protein is useful for modulating expression of a VEGF gene,
 CC thereby regulating angiogenesis and vasculogenesis. This can be used to
 CC treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
 CC diabetic retinopathy or psoriasis. The present sequence is a peptide
 CC shown in the invention.
 XX Sequence 7 AA;
 SQ Query Match 100.0%; Score 35; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRSNLTR 7
 Db |||||||
 1 DRSNLTR 7
 RESULT 13
 ABB80798
 ID ABB80798 standard; peptide; 7 AA.
 XX ABB80798;
 AC 23-SEP-2002 (first entry)
 DT Human ER-alpha locus targeting ZFP1 peptide #8.
 DE ZFP; cytostatic; antidiabetic; ophthalmological; vasotropic; chromatin;
 KW gene expression; antirheumatic; antiarthritic; antipsoriatic; nootropic;
 KW neuroprotective; cerebroprotective; estrogen receptor alpha; ER-alpha;
 KW zinc finger protein.
 XX Synthetic.
 OS WO200244386-A2.
 PN 06-JUN-2002.
 XX 30-NOV-2001; 2001WO-US45098.
 PF 01-DEC-2000; 2000US-250804P.
 PR (SANG-) SANGAMO BIOSCIENCES INC.
 PA Wolffe AP, Tse C, Collingwood T;
 XX WPI; 2002-537455/57.
 DR Regulating expression of gene by contacting cell with regulatory
 PT molecule comprising DNA-binding domain targeted to sequence within

PT accessible region of cellular chromatin associated with a gene, and
XX functional domain
PS Example 1; Page 44; 64pp; English.
XX
CC The invention relates to regulating the expression of a gene residing in
CC the chromatin of a cell. The method involves identifying one or more
CC accessible regions in cellular chromatin associated with gene; designing
CC a regulatory molecule, where the regulatory molecule comprises a DNA-
CC binding domain targeted to a sequence within the accessible region, and a
CC functional domain; and contacting the regulatory molecule with the cell.
CC The method is used for regulating the expression of a gene (e.g., a gene
CC encoding a nuclear receptor such as estrogen receptor alpha (ERalpha),
CC estrogen receptor beta (ERbeta), hepatocyte nuclear factor 4 alpha
CC (HNF4alpha), hepatocyte nuclear factor 4 gamma (HNF4gamma), peroxisome
CC proliferator activated receptor gamma (PPARGgamma), retinoid X receptor
CC alpha (RXRalpha), or constitutively active receptor alpha (CARalpha))
CC residing in the chromatin of a cell. Regulation of gene expression (such
CC as nuclear receptor genes) will be useful in treatment of various
CC diseases, including cancer, diabetes and cardiovascular disease, where
CC the regulatory molecule as described above, is contacted with the cell to
CC carry out the regulation. The method is also useful for modulation of
CC gene expression for therapeutic or prophylactic applications e.g.,
CC diabetic retinopathy, ischaemia, macular degeneration, rheumatoid
CC arthritis, psoriasis, HIV infection, sickle cell anemia, Alzheimer's
CC disease, stroke, etc. The method also has applications in pharmaceutical
CC research of both nuclear receptors of known function as well as those of
CC unknown function. The method also facilitates development of tissue and
CC animal models of disease states, drug validation, and therapeutic product
CC development. The methods also allow identification of the role of nuclear
CC receptors of unknown functions in cellular homeostasis. Sequences
CC ABB80791-817 represent zinc finger protein (ZFP) DNA-binding domains that
CC were fused to functional domains and tested for their ability to regulate
CC expression of the ER in living cells.

XX Sequence 7 AA;

Query Match 100.0%; Score 35; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSLNLR 7
DB 1 DRSLNLR 7
|||||||

RESULT 14

ABB80812 ID ABB80812 standard; peptide; 7 AA.

XX ABB80812;

XX 23-SEP-2002 (first entry)

XX Human ER-alpha locus targeting ZFP3 peptide #4.

DE ZFP: cytostatic; antidiabetic; ophthalmological; vasotropic; chromatin;
KW gene expression; antirheumatic; antiarthritic; antipsoriatic; neutropic;
KW neuroprotective; cerebroprotective; estrogen receptor alpha; ER-alpha;
KW zinc finger protein.

XX Synthetic.

XX WO200244386-A2.

XX 06-JUN-2002.

XX 30-NOV-2001; 2001WO-US45098.

XX 01-DEC-2000; 2000US-250804P.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX

PI Wolfe AP, Tse C, Collingwood T;

XX WPI; 2002-537455/57.

XX Regulating expression of gene by contacting cell with regulatory
PT molecule comprising DNA-binding domain targeted to sequence within
PT accessible region of cellular chromatin associated with a gene, and
PT functional domain

XX Example 1; Page 44; 64pp; English.

XX The invention relates to regulating the expression of a gene residing in
CC the chromatin of a cell. The method involves identifying one or more
CC accessible regions in cellular chromatin associated with gene; designing
CC a regulatory molecule, where the regulatory molecule comprises a DNA-
CC binding domain targeted to a sequence within the accessible region, and a
CC functional domain; and contacting the regulatory molecule with the cell.
CC The method is used for regulating the expression of a gene (e.g., a gene
CC encoding a nuclear receptor such as estrogen receptor alpha (ERalpha),
CC estrogen receptor beta (ERbeta), hepatocyte nuclear factor 4 alpha
CC (HNF4alpha), hepatocyte nuclear factor 4 gamma (HNF4gamma), peroxisome
CC proliferator activated receptor gamma (PPARGgamma), retinoid X receptor
CC alpha (RXRalpha), or constitutively active receptor alpha (CARalpha))
CC residing in the chromatin of a cell. Regulation of gene expression (such
CC as nuclear receptor genes) will be useful in treatment of various
CC diseases, including cancer, diabetes and cardiovascular disease, where
CC the regulatory molecule as described above, is contacted with the cell to
CC carry out the regulation. The method is also useful for modulation of
CC gene expression for therapeutic or prophylactic applications e.g.,
CC diabetic retinopathy, ischaemia, macular degeneration, rheumatoid
CC arthritis, psoriasis, HIV infection, sickle cell anemia, Alzheimer's
CC disease, stroke, etc. The method also has applications in pharmaceutical
CC research of both nuclear receptors of known function as well as those of
CC unknown function. The method also facilitates development of tissue and
CC animal models of disease states, drug validation, and therapeutic product
CC development. The methods also allow identification of the role of nuclear
CC receptors of unknown functions in cellular homeostasis. Sequences
CC ABB80791-817 represent zinc finger protein (ZFP) DNA-binding domains that
CC were fused to functional domains and tested for their ability to regulate
CC expression of the ER in living cells.

XX Sequence 7 AA;

Query Match 100.0%; Score 35; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSLNLR 7
DB 1 DRSLNLR 7
|||||||

RESULT 15

ABP48386 ID ABP48386 standard; Peptide; 7 AA.

XX ABP48386;

XX 28-AUG-2002 (first entry)

XX Zinc finger protein related peptide motif SEQ ID NO:395.

XX Zinc finger protein; ZFP; DNA binding protein; zinc finger.

XX Homo sapiens.

OS Synthetic.

XX WO200242459-A2.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US43438.

XX

```

PR 20-NOV-2000; 2000US-0716637.
XX (SANG-) SANGAMO BIOSCIENCES INC.
PA
XX Liu Q;
PI
XX WPI; 2002-500284/53.
XX
XX New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus
XX
XX Example 1; Page 37; 81pp; English.
XX
XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (I) a polypeptide
CC (II) comprising (1); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such
CC that it binds to the S2 target subsite, and selecting the F3 zinc
CC finger such that it binds to the S3 target subsite, thus designing (I)
CC that binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determine the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX SQ Sequence 7 AA;

```

```

Query Match 100.0%; Score 35; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 DRSNLTR 7
Db 1 DRSNLTR 7

```

```

Search completed: July 15, 2003, 12:53:00
Job time : 33.6667 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:50:51 ; Search time 12.3333 Seconds
(without alignments)
54.563 Million cell updates/sec

Title: US-10-006-069A-55

Perfect score: 35

Sequence: 1 DRSNLTR 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: p1r1.*

2: p1r2.*

3: p1r3.*

4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	91.4	530	2 E82491	sigma-54 dependent
2	30	85.7	125	2 H81997	probable lipoprote
3	30	85.7	140	2 S06584	finger protein (cl
4	30	85.7	377	2 A53044	geranylgeranyl-di
5	30	85.7	1523	2 T13953	MEGF5 protein - ra
6	29	82.9	198	2 AC0940	conserved hypotet
7	29	82.9	199	2 S40863	hypothetical 21.8K
8	29	82.9	199	2 E91234	hypothetical prote
9	29	82.9	199	2 E86081	hypothetical prote
10	29	82.9	220	2 T45056	hypothetical prote
11	29	82.9	276	2 D86786	positive regulator
12	29	82.9	304	2 H87572	amidase-related pr
13	29	82.9	305	2 T36056	hypothetical prote
14	29	82.9	362	2 S70961	ribD protein - vib
15	29	82.9	372	2 G30983	GDP-D-mannose dehy
16	29	82.9	372	2 T44322	hypothetical prote
17	29	82.9	372	2 B85829	GDP-mannose dehydr
18	29	82.9	373	2 S28470	GDP-mannose 4,6-de
19	29	82.9	373	2 D64971	GDP-D-mannose dehy
20	29	82.9	373	2 B30986	GDP-D-mannose dehy
21	29	82.9	373	2 E85831	GDP-D-mannose dehy
22	29	82.9	373	2 AB0769	GDPmannose 4,6-deh
23	29	82.9	377	2 B53044	geranylgeranyl-di
24	29	82.9	381	2 D87637	hypothetical prote
25	29	82.9	533	2 T36919	hypothetical prote
26	29	82.9	540	2 T14748	hypothetical prote
27	29	82.9	1466	2 A36426	SPA2 protein - yea
28	29	82.9	2529	2 A56923	transcription fact
29	29	82.9	2578	2 A56922	transcription fact

30	28	80.0	144	2 A29101	vasopressin / neur
31	28	80.0	220	2 D71980	hypothetical prote
32	28	80.0	319	2 S57969	repB protein - Rhl
33	28	80.0	340	1 A42987	N-acetyl-gamma-glu
34	28	80.0	347	2 H84673	hypothetical prote
35	28	80.0	357	2 T41314	hypothetical repea
36	28	80.0	798	2 T00131	xylan 1,4-beta-xy
37	28	80.0	953	2 C89824	hypothetical prote
38	28	80.0	3746	1 YGFLV3	alpha-aminoadipyl-
39	28	80.0	3791	1 YGFLW8	alpha-aminoadipyl-
40	27	77.1	61	2 E39741	1x hypothetical t
41	27	77.1	115	2 D70338	hypothetical prote
42	27	77.1	166	2 S76649	hypothetical prote
43	27	77.1	226	2 H71700	hypothetical prote
44	27	77.1	306	2 T47832	hypothetical prote
45	27	77.1	334	2 A10795	probable receptor/

ALIGNMENTS

RESULT 1

E82491

sigma-54 dependent transcription regulator VCA0182 [imported] - Vibrio cholerae (stra

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: E82491

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: E82491

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-530 <HEI>

A:Cross-references: GB:AE004358; GB:AE03853; NID:g9657566; PIDN:AAF96095.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VCA0182

A:Map position: 2

Query Match

Best Local Similarity 91.4%; Score 32; DB 2; Length 530;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRSNLTR 7

Db 505 DRANLTR 511

RESULT 2

H81997

probable lipoprotein NMA0065 [imported] - Neisseria meningitidis (strain Z2491 serogr

C:Species: Neisseria meningitidis

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: H81997

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

l; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: H81997

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-125 <PAR>

A:Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB83381.1; PID:g737

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA0065

C:Superfamily: Neisseria meningitidis probable lipoprotein NMA0065

Query Match 85.7%; Score 30; DB 2; Length 125;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 6
 |||||
 Db 88 DRSNLT 93

RESULT 3
 S06584
 finger protein (clone XlcGP71.1) - African clawed frog (fragment)
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 31-Dec-1993
 C:Accession: S06584
 R: Nietfeld, W.; El-Baradi, T.; Mentzel, H.; Pieler, T.; Koester, M.; Poeting, A.; Knoech
 J. Mol. Biol. 208, 639-659, 1989
 A:Title: Second-order repeats in Xenopus laevis finger proteins.
 A:Reference number: S05632; MUID:90040698; PMID:2509712
 A:Accession: S06584
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-140 <NIE>
 C:Keywords: DNA binding; zinc finger

Query Match 85.7%; Score 30; DB 2; Length 140;
 Best Local Similarity 85.7%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 7
 |||||
 Db 17 DRSNLT 23

RESULT 4
 A53044
 geranylgeranyl-diphosphate geranylgeranyltransferase (EC 2.5.1.32) I beta chain - human
 C:Species: Homo sapiens (man)
 C:Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 24-Nov-1999
 C:Accession: A53044
 R: Zhang, F.L.; Diehl, R.E.; Kohl, N.E.; Gibbs, J.B.; Giros, B.; Casey, P.J.; Omer, C.A.
 J. Biol. Chem. 269, 3175-3180, 1994
 A:Title: cDNA cloning and expression of rat and human protein geranylgeranyltransferase
 A:Reference number: A53044; MUID:94148804; PMID:8106351
 A:Accession: A53044
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-377 <ZHA>
 A:Cross-references: GB:L25441; NID:9466490; PIDN:AAA35888.1; PID:9466491
 C:Genetics:
 A:Gene: GDB:PGGT1B; GGT1; BGGI
 A:Cross-references: GDB:305477
 C:Superfamily: cell division control protein CDC43
 C:Keywords: transferase

Query Match 85.7%; Score 30; DB 2; Length 377;
 Best Local Similarity 85.7%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLT 7
 |||||
 Db 86 DRSNLT 92

RESULT 5
 T13953
 MEGF5 protein - rat
 N:Alternate names: slit protein homolog
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002
 C:Accession: T13953
 R: Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
 Genomics 51, 27-34, 1998

A:Title: Identification of high-molecular-weight proteins with multiple EGF-like moti
 A:Reference number: Z14126; MUID:98360089; PMID:9693030
 A:Accession: T13953
 A:Status: preliminary; translated from GB/EMBL/DDBT
 A:Molecule type: mRNA
 A:Residues: 1-1523 <NAK>
 A:Cross-references: EMBL:AB011531; NID:93449291; PIDN:BAA32461.1; PID:93449292
 C:Genetics:
 A:Gene: MEGF5
 C:Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein

Query Match 85.7%; Score 30; DB 2; Length 1523;
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 7
 ||::||
 Db 69 DRNNITR 75

RESULT 6
 AC0940
 conserved hypothetical protein STY3788 [imported] - Salmonella enterica subsp. enteri
 C:Species: Salmonella enterica subsp. enterica serovar typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AC0940
 R: Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
 , S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AC0940
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-138 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD09541.1; PID:g16504657; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY3788
 C:Superfamily: Escherichia coli hypothetical protein y110

Query Match 82.9%; Score 29; DB 2; Length 198;
 Best Local Similarity 71.4%; Pred. No. 32;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 7
 |::|||
 Db 73 DKANLTR 79

RESULT 7
 S40863
 hypothetical 21.8K protein (tpla 3'region) precursor - Escherichia coli (strain K-12)
 N:Alternate names: hypothetical protein f199
 C:Species: Escherichia coli
 C:Date: 19-May-1994 #sequence_revision 01-Sep-1995 #text_change 01-Mar-2002
 C:Accession: S40863; G5198
 R: Plunkett III, G.; Burland, V.; Daniels, D.L.; Blattner, F.R.
 Nucleic Acids Res. 21, 3391-3398, 1993
 A:Title: Analysis of the Escherichia coli genome. III. DNA sequence of the region fro
 A:Reference number: S40802; MUID:93347969; PMID:8346018
 A:Accession: S40863
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-199 <PLU>
 A:Cross-references: EMBL:L19201; NID:g304961; PIDN:AAB03052.1; PID:g305023
 R: Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
 .A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12:

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: C65198

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-199 <BLAT>

A:Cross-references: GB:AE000466; GB:U00096; NID:g2367328; PIDN:AAC76902.1; PID:g1790354;

A:Experimental source: strain K-12, substrain M6165

C:Genetics:

A:Gene: yliQ

C:Superfamily: Escherichia coli hypothetical protein yliQ

Query Match 82.9%; Score 29; DB 2; Length 199;

Best Local Similarity 71.4%; Pred. No. 32;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 7

I:||||

Db 73 DKANLT 79

RESULT 8

hypothetical protein ECs4845 [imported] - Escherichia coli (strain O157:H7, substrain R1

E91234

C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: E91234

R:Havashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: E91234

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-199 <HAY>

A:Cross-references: GB:BA000007; PIDN:BAH38268.1; PID:g1364321; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain R1MD 0509952

C:Genetics:

A:Gene: ECs4845

C:Superfamily: Escherichia coli hypothetical protein yliQ

Query Match 82.9%; Score 29; DB 2; Length 199;

Best Local Similarity 71.4%; Pred. No. 32;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 7

I:||||

Db 73 DKANLT 79

RESULT 9

hypothetical protein yliQ [imported] - Escherichia coli (strain O157:H7, substrain EDL93

E86081

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: E86081

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grobeck, F.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: E86081

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-199 <STO>

A:Cross-references: GB:AE005174; NID:gl2518821; PIDN:AAAG59113.1; GSPDB:GN00145; UWGP:Z54

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: yliQ

C:Superfamily: Escherichia coli hypothetical protein yliQ

Query Match 82.9%; Score 29; DB 2; Length 199;

Best Local Similarity 71.4%; Pred. No. 32;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 7

I:||||

Db 73 DKANLT 79

RESULT 10

hypothetical protein Y39B6B.dd [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Aug-2000

C:Accession: T45056

R:Wilson, R.; Ainscough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burto

raser, A.; Fulton, L.; Gardner, A.; Green, P.; Hawkins, T.; Hallier, L.; Jier, M.; Jo

B.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifken, L.; Roopra, A.; Saunders, D.

Nature 368, 32-38, 1994

A:Authors: Showkhen, R.; Sims, M.; Smalton, N.; Smith, A.; Smith, M.; Sonhammer, E.

tock, L.; Wilkinson-Sproat, J.; Wohlman, P.

A:Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.

A:Reference number: S43531; MUID:94150718; PMID:7906398

A:Accession: T45056

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-220 <WIL>

A:Cross-references: EMBL:AL132896; NID:g6434440; PIDN:CAB60935.1; PID:g6434470

A:Experimental source: clone Y39B6B

C:Genetics:

A:Map position: 3

A:Introns: 24/3; 71/2; 105/3; 140/1; 173/3; 208/2

A:Note: Y39B6B.dd

C:Superfamily: Caenorhabditis elegans hypothetical protein Y39B6B.dd

Query Match 82.9%; Score 29; DB 2; Length 220;

Best Local Similarity 71.4%; Pred. No. 36;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 7

I:||||

Db 44 DKSNT 50

RESULT 11

D86786

positive regulator [imported] - Lactococcus lactis subsp. lactis (strain ILL403)

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: D86786

R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Eh

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: D86786

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-276 <STO>

A:Cross-references: GB:AE005176; PID:g12724269; PIDN:AAK05390.1; GSPDB:GN00146

A:Experimental source: strain ILL403

C:Genetics:

A:Gene: gadR

Query Match 82.9%; Score 29; DB 2; Length 276;

Best Local Similarity 71.4%; Pred. No. 46;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 7

I:||||

Db 29 DRSNLT 35

RESULT 12

H87572

amidase-related protein [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: H87572
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:11259647
A:Accession: H87572
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-304 <STO>
A:Cross-references: GB:AE005673; NID:gl3424186; PIDN:AAK24580.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2611

Query Match 82.9%; Score 29; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RSNLTR 7
|||||
DB 264 RSNLTR 269

RESULT 13

T36056
hypothetical protein SCD78.17c - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Feb-2000
C:Accession: T36056
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z21595
A:Accession: T36056
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-305 <SAU>
A:Cross-references: EMBL:AL034355; PIDN:CAA22222.1; GSPDB:GN00070; SCOEDB:SCD78.17c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCD78.17c
C:Superfamily: Methanobacterium thermoautotrophicum conserved hypothetical protein MTH0

Query Match 82.9%; Score 29; DB 2; Length 305;
Best Local Similarity 71.4%; Pred. No. 51;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
|:|:|
DB 208 DRANVTR 214

RESULT 14

S70961
rfbD protein - Vibrio cholerae (fragment)
C:Species: Vibrio cholerae
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 20-Jun-2000
C:Accession: S70961
R:Bik, E.M.; Bunschoten, A.E.; Willems, R.J.L.; Chang, A.C.Y.; Mooi, F.R.
Mol. Microbiol. 20, 799-811, 1996
A:Title: Genetic organization and functional analysis of the otn DNA essential for cell-
A:Reference number: S70952; MUID:96386047; PMID:8793876
A:Accession: S70961
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-362 <BIK>
A:Cross-references: EMBL:X90547; NID:gl469276; PIDN:CAA62143.1; PID:gl107926
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
C:Genetics:
A:Gene: rfbD
C:Superfamily: GDP-D-mannose dehydratase

Query Match 82.9%; Score 29; DB 2; Length 362;
Best Local Similarity 85.7%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
|:|:|:|
DB 67 DSSNLTR 73

RESULT 15

G90983
GDP-D-mannose dehydratase [imported] - Escherichia coli (strain O157:H7, substrain RI
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: G90983
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: G90983
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-372 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA036262.1; PID:gl3362307; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECS2839
C:Superfamily: GDP-D-mannose dehydratase

Query Match 82.9%; Score 29; DB 2; Length 372;
Best Local Similarity 85.7%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
|:|:|:|
DB 67 DSSNLTR 73

Search completed: July 15, 2003, 12:55:29
Job time : 13.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:49:31 ; Search time 6.66667 Seconds
(without alignments)
43.550 Million cell updates/sec

Title: US-10-006-069A-55
Perfect score: 35
Sequence: 1 DRSNLTR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	DB ID	Description
1	31	88.6	276	1	ICE6_MOUSE
2	31	88.6	293	1	ICE6_HUMAN
3	30	85.7	140	1	PGT1_XENLA
4	30	85.7	377	1	PGT1_HUMAN
5	30	85.7	722	1	VGLH_GPCMV
6	29	82.9	199	1	YIIQ_ECOLI
7	29	82.9	372	1	GMD2_VIBCH
8	29	82.9	373	1	GMD4_ECOLI
9	29	82.9	373	1	GMD1_VIBCH
10	29	82.9	377	1	PGT1_RAT
11	29	82.9	379	1	CYB_ARTCI
12	29	82.9	810	1	233A_HUMAN
13	29	82.9	1466	1	SPA2_YEAST
14	28	80.0	144	1	NEU2_CAVPO
15	28	80.0	340	1	ARGC_STRCL
16	28	80.0	3746	1	ACVS_PENCH
17	28	80.0	3791	1	ACVT_PENCH
18	27	77.1	61	1	CCMD_BRAJA
19	27	77.1	77	1	RECA_SPIME
20	27	77.1	115	1	Y420_AQUAE
21	27	77.1	147	1	FMOD_RABIT
22	27	77.1	147	1	LUM_PIG
23	27	77.1	166	1	Y516_SYNY3
24	27	77.1	226	1	UPPS_RICPR
25	27	77.1	353	1	PROB_THEMA
26	27	77.1	375	1	FMOD_BOVIN
27	27	77.1	376	1	FMOD_HUMAN
28	27	77.1	376	1	FMOD_MOUSE
29	27	77.1	376	1	FMOD_RAT
30	27	77.1	860	1	BGL1_ASFAC
31	27	77.1	982	1	CELB_HUMAN
32	27	77.1	1011	1	UBAL_HUMAN
33	27	77.1	3027	1	POLG_PYFV1

34	26	74.3	104	1	PK21_GUTHI
35	26	74.3	164	1	NEU2_HUMAN
36	26	74.3	236	1	YP2C_STRAU
37	26	74.3	257	1	GPGL_SCHPO
38	26	74.3	273	1	VNST_UUK
39	26	74.3	332	1	YC56_PASMO
40	26	74.3	332	1	YIAK_HAEIN
41	26	74.3	335	1	ETV2_MOUSE
42	26	74.3	345	1	HEMS_YEREN
43	26	74.3	345	1	HMUS_YERPE
44	26	74.3	348	1	TRPD_SYNT3
45	26	74.3	373	1	ICEB_MOUSE

ALIGNMENTS

RESULT 1

ICE6_MOUSE

ID ICE6_MOUSE STANDARD; PRT; 276 AA.

AC O08738:

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Caspase-6 precursor (EC 3.4.22.-) (Apoptotic protease Mch-2).

GN CASP6.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C3H/An;

RX MEDLINE=97190206; PubMed=9036361;

RA van de Craen M., Vandenabeele P., Declercq W., van den Brande I., van Loo G., Molenans F., Schotte P., van Crieke W., Beyaert R., Fiers W.;

RT "Characterization of seven murine caspase family members.";

RL FEBS Lett. 403:61-69(1997).

CC -1- FUNCTION: INVOLVED IN THE ACTIVATION CASCADE OF CASPASES RESPONSIBLE FOR APOPTOSIS EXECUTION. CLEAVES POLY(ADP-RIBOSE) POLYMERASE IN VITRO, AS WELL AS LAMINS. OVEREXPRESSION PROMOTES PROGRAMMED CELL DEATH (BY SIMILARITY).

CC -1- SUBUNIT: HETERODIMER OF A 18 kDa (P18) AND A 11 kDa (P11) SUBUNIT (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LUNG, LIVER, KIDNEY, TESTIS, AND HEART. LOWER LEVELS IN SPLEEN, SKELETAL MUSCLE, AND BRAIN.

CC -1- PTM: CLEAVAGES BY CPP32, CASPASE-8 OR -10 GENERATE THE TWO ACTIVE SUBUNITS (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C14.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC ENBL; Y13087; CAA73529.1; -

DR HSHP; P42574; 1PAU.

DR MEROPS; C14.005; -

DR MGD; MGI:1312921; Casp6

DR InterPro; IPR002398; ICE.

DR InterPro; IPR002138; ICE.p10.

DR InterPro; IPR001309; ICE.p20.

DR Pfam; PF00655; ICE.p10; 1.

DR Pfam; PF00656; ICE.p20; 1.

DR PRINTS; PR00376; ILIBCNZYME.

DR SMART; SM00115; CASc; 1.

DR PROSITE; PS01122; CASPASE_CYS; 1.

```
DR PROSITE; PS01121; CASPASE_HIS; 1.
DR PROSITE; PS0207; CASPASE_P10; 1.
DR PROSITE; PS0208; CASPASE_P20; 1.
KW Hydrolase; Thiol protease; Apoptosis; Zymogen.
FT PROPEP 1 5 BY SIMILARITY.
FT CHAIN 6 162 CASPASE-6 SUBUNIT P18 (BY SIMILARITY).
FT PROPEP 163 176 BY SIMILARITY.
FT CHAIN 177 276 CASPASE-6 SUBUNIT P11 (BY SIMILARITY).
FT ACT_SITE 104 104 BY SIMILARITY.
FT ACT_SITE 146 146 BY SIMILARITY.
SQ SEQUENCE 276 AA; 31595 MW; 5965DE9321126B6C CRC64;

Query Match 88.6%; Score 31; DB 1; Length 276;
Best Local Similarity 85.7%; Pred. No. 6.3;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLT 7
DB 53 DRDNLTR 59

RESULT 2
ICE6_HUMAN STANDARD; PRT; 293 AA.
AC P55212;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Caspase-6 precursor (EC 3.4.22.-) (Apoptotic protease Mch-2).
GN CASP6 OR MCH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RA Fernandes-Alnemri T., Litwack G., Alnemri E.S.;
RX "Mch2, a new member of the apoptotic Ced-3/Ice cysteine protease gene family";
RT Cancer Res. 55:2737-2742(1995).
RL [2]
RN RP
RP PROCESSING.
RX MEDLINE=97059171; PubMed=8902001;
RA Srinivasula S.M., Fernandes-Alnemri T., Zangrilli J., Robertson N.,
RA Armstrong R.C., Wang L., Trapani J.A., Tomaselli K.J., Litwack G.,
RA Alnemri E.S.;
RT "The Ced-3/interleukin 1beta converting enzyme-like homolog Mch6 and the lamin-cleaving enzyme Mch2alpha are substrates for the apoptotic mediator CPP32.";
RL J. Biol. Chem. 271:27099-27106(1996).
CC -!- FUNCTION: INVOLVED IN THE ACTIVATION CASCADE OF CASPASES
CC RESPONSIBLE FOR APOPTOSIS EXECUTION. CLEAVES POLY(ADP-RIBOSE)
CC POLYMERASE IN VITRO, AS WELL AS LAMINS. OVEREXPRESSION PROMOTES
CC PROGRAMMED CELL DEATH.
CC -!- SUBUNIT: HETEROOLIGOMER OF A 18 kDa (P18) AND A 11 kDa (P11) SUBUNIT.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA (SHOWN HERE) AND BETA;
CC ARE PRODUCED BY ALTERNATIVE SPLICING. THE BETA ISOFORM DOES NOT
CC SEEM TO HAVE PROTEOLYTIC ACTIVITY.
CC -!- PTM: CLEAVAGES BY CPP32, CASPASE-8 OR -10 GENERATE THE TWO ACTIVE
CC SUBUNITS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C14.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; U20536; AAC50168.1; -.
DR EMBL; U20537; AAC50169.1; -.
DR HSSP; P42574; 1PAU.
DR MEROPS; C14.005; -.
DR GENNEW; HGNC:1507; CASP6.
DR MIM; 601532; -.
DR InterPro; IPR002398; ICE.
DR InterPro; IPR002138; ICE_P10.
DR InterPro; IPR001309; ICE_P20.
DR Pfam; PF00655; ICE_P10; 1.
DR Pfam; PF00656; ICE_P20; 1.
DR PRINTS; PR00376; ILIBENZYM.
DR SMART; SM00115; CASC; 1.
DR PROSITE; PS01122; CASPASE_CYS; 1.
DR PROSITE; PS01121; CASPASE_HIS; 1.
DR PROSITE; PS0207; CASPASE_P10; 1.
DR PROSITE; PS0208; CASPASE_P20; 1.
KW Hydrolase; Thiol protease; Apoptosis; Zymogen; Alternative splicing.
FT PROPEP 1 23
FT CHAIN 24 179 CASPASE-6 SUBUNIT P18.
FT PROPEP 180 193
FT CHAIN 194 293 CASPASE-6 SUBUNIT P11.
FT ACT_SITE 121 121 BY SIMILARITY.
FT ACT_SITE 163 163 BY SIMILARITY.
FT VARSPIC 14 102 MISSING (IN ISOFORM BETA).
SQ SEQUENCE 293 AA; 33409 MW; BD9204E23CE1F670 CRC64;

Query Match 88.6%; Score 31; DB 1; Length 293;
Best Local Similarity 85.7%; Pred. No. 6.7;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLT 7
DB 70 DRDNLTR 76

RESULT 3
ZG7L_XENLA STANDARD; PRT; 140 AA.
ID ZG7L_XENLA
AC P18736;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Gastrula zinc finger protein XLCGF71.1 (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90040698; PubMed=2509712;
RA Nietfeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,
RA Poeling A., Knoechel W.;
RT "Second-order repeats in Xenopus laevis finger proteins.";
RL J. Mol. Biol. 208:639-659(1989).
DR PIR; S06584; S06584.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 5.
DR SMART; SM00355; Znf_C2H2; 5.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 5.
KW Zinc-finger; Metal-binding; DNA-binding; Repeat.
FT NON_TER 1
FT ZN_FING 6 28 C2H2-TYPE.
FT ZN_FING 34 56 C2H2-TYPE.
FT ZN_FING 62 84 C2H2-TYPE.
FT ZN_FING 90 112 C2H2-TYPE.
FT ZN_FING 118 140 C2H2-TYPE.
FT NON_TER 140 140
SQ SEQUENCE 140 AA; 16098 MW; 3D37A0D462B664CE CRC64;
```


Query Match 85.7%; Score 30; DB 1; Length 140;
Best Local Similarity 85.7%; Pred. No. 4.9;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRSLNLR 7
|||||
Db 17 DRSLNLR 23

RESULT 4
PGLT_HUMAN
ID PGLT_HUMAN STANDARD; PRT; 377 AA.
AC P53609;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Geranylgeranyl transferase type I beta subunit (EC 2.5.1.-) (Type I
protein geranyl-geranyltransferase beta subunit) (GGrase-I-beta).
GN PGRIB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUP-Placenta, and Kidney;
RX MEDLINE=94148804; PubMed=8106351;
RA Zhang F.L., Diehl R.E., Kohl N.E., Gibbs J.B., Giros B.,
Casey P.J., Omer C.A.;
RT "cDNA cloning and expression of rat and human protein
geranylgeranyltransferase type-I.";
RL J. Biol. Chem. 269:3175-3180(1994).
CC -1- FUNCTION: CATALYZES THE TRANSFER OF A GERANYL-GERANYL MOIETY FROM
GERANYLGERANYL PYROPHOSPHATE TO PROTEINS HAVING THE C-TERMINAL
SEQUENCE CYS-ALIPHATIC-ALIPHATIC-X. ACTS ON THE RAC1, RAC2,
RAP1A AND RAP1B PROTEINS.
CC -1- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC -1- SIMILARITY: BELONGS TO THE PROTEIN PRENYLTRANSFERASE BETA SUBUNIT
FAMILY.
CC -1- SIMILARITY: CONTAINS 4 PFTB REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; L25441; AAA35888.1; -
CC Genew; HGNC:8895; PGGT1B.
CC MIM: 602031; -
CC InterPro; IPR001330; Prenyltrans.
CC Pfam; PF00432; Prenyltrans; 5.
KW Transferase; Prenyltransferase; Repeat; Zinc.
FT REPEAT 144 186 PFTB 1.
FT REPEAT 193 234 PFTB 2.
FT REPEAT 245 284 PFTB 3.
FT REPEAT 291 333 PFTB 4.
FT METAL 269 269 ZINC (BY SIMILARITY).
FT METAL 271 271 ZINC (BY SIMILARITY).
FT METAL 321 321 ZINC (BY SIMILARITY).
SQ SEQUENCE 377 AA; 42396 MW; 565CD9B6C087A4DF CRC64;

Query Match 85.7%; Score 30; DB 1; Length 377;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRSLNLR 7
|||||
Db 86 DRSLNLR 92

RESULT 5
VGLH_GPCMV
ID VGLH_GPCMV STANDARD; PRT; 722 AA.
AC P87730;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein H precursor.
GN GH OR UL75.
OS Guinea pig cytomegalovirus (strain 22122 / ATCC VR682) (GPCMV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_TaxID=103920;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=98187294; PubMed=9526546;
RA Brady R.C., Schleiss M.R.;
RT "Identification and characterization of the guinea-pig
cytomegalovirus glycoprotein H gene.";
RL Arch. Virol. 141:2409-2424(1996).
CC -1- FUNCTION: IMPORTANT DETERMINANT OF VIRUS INFECTIVITY.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN H FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC

EMBL; U49361; AAC56576.1; -
DR InterPro; IPR003493; Herpes_glycop.
DR Pfam; PF02489; Herpes_glycop_H; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 ?
FT CHAIN ? 722 GLYCOPROTEIN H.
FT TRANSMEM 696 716 POTENTIAL.
FT CARBOHYD 143 143 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 542 542 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 680 680 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 722 AA; 81609 MW; 13E0424C81FCA16 CRC64;

Query Match 85.7%; Score 30; DB 1; Length 722;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRSLNLT 6
|||||
Db 539 DRSLNLT 544

RESULT 6
YIIQ_ECOLI
ID YIIQ_ECOLI STANDARD; PRT; 199 AA.
AC P32160;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yIIQ precursor.
GN YIIQ OR B3920.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-K12 / MG1655;
RX MEDLINE-93347969; PubMed-8346018;
RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the
RL region from 87.2 to 89.2 minutes.";
RL Nucleic Acids Res. 21:3391-3398(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L19201; AAB03052.1;
DR EMBL; AE000466; AAC76902.1;
DR FIR; S40863; S40863.
DR EcoGene; Egl1874; yliQ.
KW Hypothetical protein; signal; Complete proteome.
FT SIGNAL 1 23 POTENTIAL
FT CHAIN 24 199 HYPOTHETICAL PROTEIN YIIQ.
SQ SEQUENCE 199 AA; 21763 MW; C946DA941973F334 CRC64;
-----
Query Match 82.9%; Score 29; DB 1; Length 199;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
*QY 1 DRSNLTR 7
DB 73 DRANLTR 79
-----
RESULT 7
ID GMD2_VIBCH STANDARD; PRT; 372 AA.
AC Q56598;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable GDP-mannose 4,6-dehydratase (EC 4.2.1.47) (GDP-D-mannose
DE dehydratase).
GN rFBD OR GMD.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AI-1837 / Serotype O139;
RX MEDLINE-97252505; PubMed-9098074;
RA Stroeder U.H., Parasivam G., Dredge B.K., Manning P.A.;
RT "Novel Vibrio cholerae O139 genes involved in lipopolysaccharide
RT biosynthesis.";
RL J. Bacteriol. 179:2740-2747(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MO45 / ATCC 51394 / Serotype O139;
RX MEDLINE-99453293; PubMed-10521656;
RA Yamasaki S., Shimizu T., Hoshino K., Ho S.-T., Shimada T., Nair G.B.,
RA Takeda Y.;
RT "The genes responsible for O-antigen synthesis of Vibrio cholerae O139
RT are closely related to those of Vibrio cholerae O22.";
RL Gene 237:321-332(1999).
RN [3]
RP SEQUENCE OF 1-348 FROM N.A.
RC STRAIN-MO45 / ATCC 51394 / Serotype O139;
RA Dumontier S.E., Escuyer V.E., Berche P.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: GDP-mannose -> GDP-4-dehydro-6-deoxy-D-mannose
CC + H(2)O.
CC -1- COFACTOR: NAD(+).
CC -1- PATHWAY: Conversion of GDP-mannose to GDP-fucose; first step.
CC -1- SIMILARITY: BELONGS TO THE GDP-MANNOSE 4,6-DEHYDRATASE FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y07786; CAA69111.1;
DR EMBL; AB012956; BAA33595.1;
DR EMBL; U24571; AAA7032.1;
DR HSSP; P32054; 1DB3.
DR InterPro; IPR001509; Epimerase_Dh.
DR Pfam; PF01370; Epimerase; 1
KW Lipopolysaccharide biosynthesis; Lyase; NAD.
SQ SEQUENCE 372 AA; 41992 MW; 11762663764B3E57 CRC64;
-----
Query Match 82.9%; Score 29; DB 1; Length 372;
Best Local Similarity 85.7%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
*QY 1 DRSNLTR 7
DB 67 DSSNLTR 73
-----
RESULT 8
ID GMD4_ECOLI STANDARD; PRT; 373 AA.
AC P32054; P77687;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GDP-mannose 4,6-dehydratase (EC 4.2.1.47) (GDP-D-mannose dehydratase).
GN GMD OR B2053 OR P3217 OR ECS2858.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE-96326333; PubMed-8759852;
RA Stevenson G., Andrianopoulos K., Hobbs M., Reeves P.R.;
RT "Organization of the Escherichia coli K-12 gene cluster responsible
RT for production of the extracellular polysaccharide colanic acid.";
RL J. Bacteriol. 178:4885-4893(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1232-1243(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE-97251358; PubMed-9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sampei G., Seki Y., Sivasubramanian S., Tagami H.,
RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
RN [4]

```

RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kubara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22(2001).
RN [6]
RP SEQUENCE OF 50-373 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=95115532; PubMed=7815923;
RA Aoyama K., Haase A.M., Reeves P.R.;
RT "Evidence for effect of random genetic drift on G+C content after
lateral transfer of fucose pathway genes to Escherichia coli K-12";
RL Mol. Biol. Evol. 11:829-838(1994).
RN [7]
RP CHARACTERIZATION.
RX MEDLINE=97400210; PubMed=9257704;
RA Sturla L., Bisso A., Zanardi D., Benatti U., de Flora A., Tonetti M.;
RT "Expression, purification and characterization of GDP-D-mannose 4,6-
dehydratase from Escherichia coli";
RL FEBS Lett. 412:126-130(1997).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=20139699; PubMed=10673432;
RA Sonoza J.R., Menon S.S., Schmidt H., Joseph-McCarthy D., Dessen A.,
RA Stahl M.L., Somers W.S., Sullivan F.X.;
RT "Structural and kinetic analysis of Escherichia coli GDP-mannose 4,6
dehydratase provides insights into the enzyme's catalytic mechanism
and regulation by GDP-fucose";
RL Structure 8:123-135(2000).
CC -!- CATALYTIC ACTIVITY: GDP-mannose = GDP-4-dehydro-6-deoxy-D-mannose
+ H(2)O.
CC -!- COFACTOR: NADP(+).
CC -!- PATHWAY: INVOLVED IN THE BIOSYNTHESIS OF THE SLIME POLYSACCHARIDE
COLANIC ACID. FIRST OF THE THREE STEPS IN THE BIOSYNTHESIS OF GDP-
FUCCOSE FROM GDP-MANNOSE.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SIMILARITY: BELONGS TO THE GDP-MANNOSE 4,6-DEHYDRATASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U38473; AAC77842.1; -;
CC EMBL: AE000295; AAC75114.1; -;
CC EMBL: D90843; BAA15909.1; -;
CC EMBL: AE005431; AAG57113.1; -;
CC EMBL: AP002559; BAB36281.1; -;
CC PDB: 1DB3; 24-NOV-99.
CC EcoGene: EG11787; gmd.
CC InterPro: IPR001509; Epimerase_Dh.
CC Pfam: PF01370; Epimerase; 1.
CC Lipopolysaccharide biosynthesis; Lyase; NADP; 3D-structure;
KW

KW Complete proteome. 135 BASE.
FT ACT_SITE 135
SQ SEQUENCE 373 AA; 42047 MW; 1A9BA2A7C566DE11 CRC64;

Query Match 82.9%; Score 29; DB 1; Length 373;
Best Local Similarity 85.7%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
Db 67 DTSNLTR 73

RESULT 9
GMDL_VIBCH STANDARD; PRT; 373 AA.
ID GMDL_VIBCH
AC Q06952;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable GDP-mannose 4,6-dehydratase (EC 4.2.1.47) (GDP-D-mannose
DE dehydratase).
GN RFED OR GMD OR VC0243.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor O17 / Serotype O1;
RX MEDLINE=92212870; PubMed=1372980;
RA Stroehner U.H., Karageorgos L.E., Morona R., Manning P.A.;
RT "Serotype conversion in Vibrio cholerae O1";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2566-2570(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor O17 / Serotype O1;
RX MEDLINE=95309704; PubMed=7540582;
RA Manning P.A., Stroehner U.H., Karageorgos L.E., Morona R.;
RT "Putative O-antigen transport genes within the rib region of Vibrio
cholerae O1 are homologous to those for capsule transport.";
RL Gene 158:1-7(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae";
RL Nature 406:477-483(2000).
CC -!- CATALYTIC ACTIVITY: GDP-mannose = GDP-4-dehydro-6-deoxy-D-mannose
+ H(2)O.
CC -!- COFACTOR: NAD(+). (BY SIMILARITY).
CC -!- PATHWAY: Conversion of GDP-mannose to GDP-fucose; first step.
CC -!- SIMILARITY: BELONGS TO THE GDP-MANNOSE 4,6-DEHYDRATASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X59554; CAA42136.1; -;
CC EMBL: AE004113; AAF93419.1; -;
CC HSP: P32054; 1DB3.
CC TIGR: VC0243; -;

DR InterPro; IPR001509; Epimerase_Dh.
 DR Pfam; PF01370; Epimerase; 1.
 KW Lipopolysaccharide biosynthesis; Lysase; NAD; Complete proteome.
 SQ SEQUENCE 373 AA; 42053 MW; 4EB10DCAB6A8C1D2 CRC64;

Query Match 82.9%; Score 29; DB 1; Length 373;
 Best Local Similarity 85.7%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLTNR 7
 DB 68 DSSNLTNR 74

RESULT 10
 ID PCT1_RAT STANDARD; PRT; 377 AA.
 AC P53610;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Geranylgeranyl transferase type I beta subunit (EC 2.5.1.1*) (Type I
 DE protein geranyl-geranyltransferase beta subunit) (GGTase-I-beta).
 GN PGGTIB.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=94148804; PubMed=8106351;
 RA Zhang F.L., Diehl R.E., Kohl N.E., Gibbs J.B., Giros B.,
 RA Casey P.J., Omer C.A.;
 RT "cDNA cloning and expression of rat and human protein
 RT geranylgeranyltransferase type-I";
 RL J. Biol. Chem. 269:3175-3180(1994).
 CC -1- FUNCTION: CATALYZES THE TRANSFER OF A GERANYL-GERANYL MOIETY FROM
 CC GERANYLGERANYL PYROPHOSPHATE TO PROTEINS HAVING THE C-TERMINAL
 CC SEQUENCE CYS-ALIPHATIC-ALIPHATIC-X. ACTS ON THE RAC1, RAC2,
 CC RAP1A AND RAP1B PROTEINS.
 CC -1- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE PROTEIN PRENYLTRANSFERASE BETA SUBUNIT
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 4 PFTB REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L24116; AAA17756.1; -
 DR InterPro; IPR001330; Prenyltrans.
 DR Pfam; PF00432; prenyltrans; 5.
 KW Transferase; Prenyltransferase; Repeat; Zinc.
 FT REPEAT 144 186 PFTB 1.
 FT REPEAT 193 234 PFTB 2.
 FT REPEAT 245 284 PFTB 3.
 FT REPEAT 291 333 PFTB 4.
 FT METAL 269 269 ZINC (BY SIMILARITY).
 FT METAL 271 271 ZINC (BY SIMILARITY).
 FT METAL 321 321 ZINC (BY SIMILARITY).
 SQ SEQUENCE 377 AA; 42414 MW; AADEC7301A4A011 CRC64;

Query Match 82.9%; Score 29; DB 1; Length 377;
 Best Local Similarity 85.7%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLTNR 7

DB 86 DRSNLTNR 92

RESULT 11
 ID CYB_ARTCI STANDARD; PRT; 379 AA.
 AC Q95726; Q34305;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cytochrome B.
 DE Cytochrome B.
 GN MTCYB OR COB OR CYTB.
 OS Artibeus cinereus (Gervais' fruit-eating bat) (Dermanura cinerea).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Stenodermatinae; Artibeus.
 OX NCBI_TaxID=40224;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TK 18790 / AMNH 267197;
 RA den Bussche R.A., Hudgeons J.L., Baker R.J.;
 RT "Phylogenetic accuracy, stability, and congruence: relationships
 RT within and among the New World bat genera Artibeus, Dermanura, and
 RT Koopmanina";
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-134 FROM N.A.
 RC STRAIN=TK 14594; TISSUE=Muscle;
 RX MEDLINE=94018594; PubMed=8412654;
 RA den Bussche R.A., Baker R.J., Wichman H.A., Hamilton M.J.;
 RT "Molecular phylogenetics of Stenodermatina bat genera: congruence of
 RT data from nuclear and mitochondrial DNA";
 RL Mol. Biol. Evol. 10:944-959(1993).
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS.
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN.
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U66511; AAB06766.1; -
 DR EMBL; L19511; AAA67850.1; -
 DR InterPro; IPR000179; Cyt_b_b6.
 DR Pfam; PF00032; cytochrome_b_c; 1.
 DR Pfam; PF00033; cytochrome_b_n; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 DR PROSITE; PS00193; CYTOCHROME_B_OO; 1.
 KW Mitochondrion; Electron transport; Respiratory chain; Transmembrane;
 KW Heme.
 FT METAL 83 83 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 97 97 IRON 2 (HEME B566 AXIAL LIGAND).
 FT METAL 182 182 IRON 2 (HEME B562 AXIAL LIGAND).
 FT METAL 196 196 IRON 1 (HEME B566 AXIAL LIGAND).
 FT METAL 196 196 IRON 1 (HEME B566 AXIAL LIGAND).
 FT CONFLICT 14 14 I -> V (IN REF. 2).
 SQ SEQUENCE 379 AA; 42697 MW; 8E90A6FF5E2D3DE8 CRC64;

Query Match 82.9%; Score 29; DB 1; Length 379;
 Best Local Similarity 71.4%; Pred. No. 27;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 DRSNLT 7
Db 171 DKANLTR 177

RESULT 12
ID Z33A_HUMAN STANDARD; PRT; 810 AA.
AC O06730;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zinc finger protein 33A (zinc finger protein KOX31) (HA0946).
GN ZNF33A OR ZNF33 OR KOX31 OR KIAA0065.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96051398; PubMed=7584044;
RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
RA Seki N., Kawabayashi Y., Ishikawa K.-I., Tabata S.;
RA "Prediction of the coding sequences of unidentified human genes. II.
RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 1:223-229(1994).
RN (2)
RP SEQUENCE OF 629-810 FROM N.A.
RX MEDLINE=93219119; PubMed=8464732;
RA Tunnaciffe A., Liu L., Moore J.K., Leversha M.A., Jackson M.S.,
RA Ferguson-Smith M.A., Thiesen H.-J., Ponder B.A.;
RT "Duplicated KOX zinc finger gene clusters flank the centromere of
RT human chromosome 10: evidence for a pericentric inversion during
RT primate evolution.";
RL Nucleic Acids Res. 21:1409-1417(1993).
CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -1- SIMILARITY: CONTAINS 1 KRAB DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D31763; BAA06541.1; -
CC EMBL; X68687; CAA48646.1; -
CC EMBL; X68689; CAA48648.1; -
CC HSSP; P08045; 1ZNF.
CC TRANSFAC; T04984; -
CC Genew; HGNC:13096; ZNF33A.
CC InterPro; IPR001909; KRAB.
CC InterPro; IPR000822; Znf_C2H2.
CC Pfam; PF00096; zf-C2H2; 16.
CC Pfam; PF01352; KRAB; 1.
CC PRINTS; PR00048; ZINCFINGER.
CC PRODOM; PD000003; Znf_C2H2; 13.
CC SMART; SM00349; KRAB; 1.
CC SMART; SM00355; Znf_C2H2; 16.
CC PROSITE; PS50805; KRAB; 1.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 16.
CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 16.
CC Transcription regulation; Zinc-finger; DNA-binding; Metal-binding;
KW Nuclear protein; Repeat.
FT DOMAIN 12 83 KRAB.
FT ZN_FING 328 350 C2H2-TYPE.
FT ZN_FING 356 378 C2H2-TYPE.
```

```
FT 2N_FING 384 406 C2H2-TYPE.
FT 2N_FING 412 434 C2H2-TYPE.
FT 2N_FING 440 462 C2H2-TYPE.
FT 2N_FING 468 490 C2H2-TYPE.
FT 2N_FING 496 518 C2H2-TYPE.
FT 2N_FING 524 546 C2H2-TYPE.
FT 2N_FING 552 574 C2H2-TYPE.
FT 2N_FING 580 602 C2H2-TYPE.
FT 2N_FING 608 630 C2H2-TYPE.
FT 2N_FING 636 658 C2H2-TYPE.
FT 2N_FING 664 686 C2H2-TYPE.
FT 2N_FING 692 714 C2H2-TYPE.
FT 2N_FING 720 742 C2H2-TYPE.
FT 2N_FING 748 770 C2H2-TYPE.
SQ SEQUENCE 810 AA; 94384 MW; 96AD969EF541B73E CRC64;

Query Match 82.9%; Score 29; DB 1; Length 810;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 DRSNLT 7
Db 367 DKSNLTK 373

RESULT 13
SPA2_YEAST STANDARD; PRT; 1466 AA.
ID SPA2_YEAST
AC P23201;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SPA2 protein.
GN SPA2 OR PEAL OR YLL021W OR L1209.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=91009481; PubMed=2211820;
RA Gehrung S., Snyder M.;
RT "The SPA2 gene of Saccharomyces cerevisiae is important for
RT pheromone-induced morphogenesis and efficient mating.";
RL J. Cell Biol. 111:1451-1464(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RA Purnelle B., Goffeau A.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN PHEROMONE-INDUCED MORPHOGENESIS AND
CC EFFICIENT MATING, PERHAPS AS A CYTOSKELETAL PROTEIN.
CC -1- SUBCELLULAR LOCATION: LOCALIZES A SHARP PATCH AT THE SHMOO TIP
CC (MATING PROJECTION) WHICH IS THE SITE OF POLARIZED CELL GROWTH.
CC -1- SIMILARITY: TO PROTEINS THAT FORM COILED-COIL STRUCTURES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X53731; CAA37763.1; -
CC EMBL; X97560; CAA66170.1; -
CC EMBL; Z73126; CAA97469.1; -
CC PIR; A36426; A36426.
CC SGD; S0003944; SPA2.
CC Cell shape; Repeat; Coiled coil.
KW DOMAIN 286 388 COILED COIL (POTENTIAL).
FT DOMAIN 818 1087 25 X 9 AA APPROXIMATE TANDEM REPEATS.
```

```

FT REPEAT 818 825
FT REPEAT 826 834
FT REPEAT 835 843
FT REPEAT 860 868
FT REPEAT 875 883
FT REPEAT 884 892
FT REPEAT 893 901
FT REPEAT 902 910
FT REPEAT 911 919
FT REPEAT 920 928
FT REPEAT 929 937
FT REPEAT 938 946
FT REPEAT 947 953
FT REPEAT 954 961
FT REPEAT 962 970
FT REPEAT 971 979
FT REPEAT 980 988
FT REPEAT 989 997
FT REPEAT 998 1006
FT REPEAT 1007 1015
FT REPEAT 1036 1044
FT REPEAT 1045 1053
FT REPEAT 1054 1062
FT REPEAT 1072 1080
FT REPEAT 1081 1087
FT DOMAIN 1169 1189
FT DOMAIN 1275 1302
SQ SEQUENCE 1466 AA; 163142 MW; 2EBB616152382C89 CRC64;

Query Match 82.9%; Score 29; DB 1; Length 1466;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLTNR 7
DB 34 DRSNSTR 40

RESULT 14
NEU2_CAVPO STANDARD; PRT; 144 AA.
AC P10769;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE Vasopressin-neurophysin 2-copeptin precursor [Contains: Arg-
DE vasopressin; Neurophysin 2 (Neurophysin-II); Copeptin].
GN AVP.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE.
RX MEDLINE=87105929; PubMed=3803579;
RA Chauvet M.-T., Rouille Y., Chauvet J., Acher R.;
RT "Guinea pig neurohypophysial hormones. Peculiar processing of the
RT three-domain vasopressin precursor.";
RL FEBS Lett. 210:40-44(1987).
RN [2]
RP SEQUENCE OF 13-144.
RX MEDLINE=87247214; PubMed=3595848;
RA Chauvet J., Chauvet M.-T., Acher R.;
RT "Conformation limited proteolysis in the common neurophysin-copeptin
RT precursor shown by trypsin-sepharose chromatographic proteolysis.";
RL FEBS Lett. 217:180-183(1987).
RN [3]
RP SEQUENCE OF 13-105.
RX MEDLINE=88138574; PubMed=3436704;
RA Chauvet M.-T., Chauvet J., Acher R.;
RT "Guinea pig MSEL-neurophysin. Sequence comparison of eight mammalian
RT MSEL-neurophysins.";
RL Int. J. Pept. Protein Res. 30:676-682(1987).

```

```

RN [4]
RX MEDLINE=86136563; PubMed=3081370;
RA Chauvet M.-T., Chauvet J., Acher R.;
RT "Guinea pig copeptin. The glycopeptide domain of the vasopressin
RT precursor.";
RL FEBS Lett. 197:169-172(1986).
CC -1- FUNCTION: NEUROPHYSIN 2 SPECIFICALLY BINDS VASOPRESSIN.
CC -1- FUNCTION: VASOPRESSIN HAS A DIRECT ANTIDIURETIC ACTION ON THE
CC KIDNEY. IT ALSO CAUSES VASOCONSTRICTION OF THE PERIPHERAL VESSELS.
CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR PIR; A23630; A23630.
DR PIR; A29101; A29101.
DR PIR; S00009; S00009.
DR PIR; JS0300; JS0300.
DR HSSP; P01180; INPO.
DR InterPro; IPR000981; Neurohyp_horm.
DR Pfam; PF00184; hormone5; 1.
DR Pfam; PF00220; hormone4; 1.
DR PRINTS; PRO0831; NEUROPHYSIN.
DR ProDom; PD001676; Neurohyp_horm; 1.
DR SMART; SM00003; NH; 1.
DR PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Hypothalamus; Amidation; Cleavage on pair of basic residues;
KW Vasoconstrictor; Glycoprotein.
FT PEPTIDE 1 9 ARG-VASOPRESSIN.
FT PEPTIDE 13 105 NEUROPHYSIN 2.
FT PEPTIDE 107 144 COPEPTIN.
FT DISULFID 1 6
FT DISULFID 22 66 BY SIMILARITY.
FT DISULFID 25 39 BY SIMILARITY.
FT DISULFID 33 56 BY SIMILARITY.
FT DISULFID 40 46 BY SIMILARITY.
FT DISULFID 73 85 BY SIMILARITY.
FT DISULFID 79 97 BY SIMILARITY.
FT DISULFID 86 91 BY SIMILARITY.
FT MOD_RES 9 9
FT CARBOHYD 112 112
SQ SEQUENCE 144 AA; 15068 MW; CE2B18A162C9ABEA CRC64;

Query Match 80.0%; Score 28; DB 1; Length 144;
Best Local Similarity 71.4%; Pred. No. 15;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTNR 7
DB 109 DRSNVTQ 115

RESULT 15
ARGC_STRCL STANDARD; PRT; 340 AA.
AC P54896;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) (AGPR) (N-
DE acetyl-glutamate semialdehyde dehydrogenase) (NAGSA dehydrogenase).
GN ARGC.
OS Streptomyces clavuligerus.
OC Bacteria; Actinobacteria; Actinobacteriales; Streptomyces; Streptomyces.
OX NCBI_TaxID=1901;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27064 / DSM 738 / NRRL 3585;
RX MEDLINE=92325051; PubMed=1339424;
RA Ludovice M., Martin J.F., Carrachas P., Liras P.;
RT "Characterization of the Streptomyces clavuligerus argC gene encoding
RT N-acetylglutamyl-phosphate reductase: expression in Streptomyces
RT lividans and effect on clavulanic acid production.";
RL J. Bacteriol. 174:4606-4613(1992).
CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)

```

```
CC + phosphate = N-acetyl-5-glutamyl phosphate + NADPH.
CC -!- PATHWAY: Arginine biosynthesis; third step.
CC -!- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M83659; AAA26704.1; -
CC InterPro; IPR000706; AGPR_act_site.
CC InterPro; IPR000534; Semialdh_dh.
CC Pfam; PF01118; Semialdehyde_dh; 1.
CC Pfam; PF02774; Semialdehyde_dhc; 1.
CC ProDom; PD003765; AGPR_act_site; 1.
CC PROSITE; PS01224; ARGCC; 1.
CC Arginine biosynthesis; Oxidoreductase; NADP.
KW Arginine biosynthesis; ARGCC; 1.
FT ACT_SITE 145 145 BY SIMILARITY.
SQ SEQUENCE 340 AA; 35299 MW; BC270C115D9B494F CRC64;

Query Match      80.0%; Score 28; DB 1; Length 340;
Best Local Similarity 71.4%; Pred. NO. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 DRSNLTR 7
DB      303 DRDNLTK 309
```

Search completed: July 15, 2003, 12:53:28
Job time : 7.66667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:50:16 ; Search time 23.6667 seconds
(without alignments)
60.944 Million cell updates/sec

Title: US-10-006-069A-55
Perfect score: 35
Sequence: 1 DRSNLTR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rviris.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	91.4	530	16	Q9KMY4 vibrio chol
2	31	88.6	241	11	Q9D213 mus musculu
3	31	88.6	276	11	Q9D089 mus musculu
4	31	88.6	276	11	Q9D089 mus musculu
5	31	88.6	293	4	Q9BQE7 Q9bqe7 homo sapien
6	31	88.6	761	13	P79793 gallus gall
7	31	88.6	1941	13	Q8UWA0 gallus gall
8	30	85.7	125	16	Q9JXA9 neisseria m
9	30	85.7	182	2	Q9RBT4 pseudomonas
10	30	85.7	204	4	Q9NPX7 Q9np7 homo sapien
11	30	85.7	292	4	Q96G44 homo sapien
12	30	85.7	292	4	Q9BSG1 Q9bsg1 homo sapien
13	30	85.7	321	2	Q52190 enterococcu
14	30	85.7	321	2	Q52195 enterococcu
15	30	85.7	395	4	Q9UMC5 Q9umc5 homo sapien
16	30	85.7	413	2	Q8RNN1 legionella

17 30 85.7 451 4 Q95363 homo sapien
18 30 85.7 1515 13 Q9DE37 Q9de37 brachydanio
19 30 85.7 1523 4 Q75094 Q75094 homo sapien
20 30 85.7 1523 11 Q88280 Q88280 rattus norv
21 30 85.7 1523 11 Q9WVB4 Q9wvb4 mus musculu
22 29 82.9 183 13 Q9DDG0 Q9ddg0 oncorhynch
23 29 82.9 198 16 Q8ZKP6 Q8zkp6 salmonella
24 29 82.9 198 16 Q8Z2V3 Q8z2v3 salmonella
25 29 82.9 199 15 Q8X7A4 Q8x7a4 escherichia
26 29 82.9 220 5 Q9NET0 Q9net0 caenorhabdi
27 29 82.9 258 16 Q98922 Q98922 rhizobium 1
28 29 82.9 276 2 Q30416 Q30416 lactococcus
29 29 82.9 276 16 Q9CG18 Q9cg18 lactococcus
30 29 82.9 304 16 Q9A553 Q9a553 caulobacter
31 29 82.9 305 16 Q9ZBY2 Q9zby2 streptomyce
32 29 82.9 362 2 Q56659 Q56659 vibrio chol
33 29 82.9 364 8 Q957Q3 Q957q3 clethrionom
34 29 82.9 364 8 Q94NY7 Q94ny7 clethrionom
35 29 82.9 372 2 Q51839 Q51839 vibrio chol
36 29 82.9 372 2 Q93Q33 Q93q33 salmonella
37 29 82.9 372 2 Q9S522 Q9s522 escherichia
38 29 82.9 372 16 Q85339 Q85339 escherichia
39 29 82.9 373 16 Q8Z5H1 Q8z5h1 salmonella
40 29 82.9 373 16 Q9F7A4 Q9f7a4 salmonella
41 29 82.9 377 2 Q9LA94 Q9la94 aeromonas h
42 29 82.9 381 16 Q9A3R8 Q9a3r8 caulobacter
43 29 82.9 416 11 Q8R573 Q8r573 mus musculu
44 29 82.9 533 16 Q9X9W7 Q9x9w7 streptomyce
45 29 82.9 540 4 Q9UG14 Q9ug14 homo sapien

ALIGNMENTS

RESULT 1

Q9KMY4 ID Q9KMY4 PRELIMINARY; PRT; 530 AA.
AC Q9KMY4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Sigma-54 dependent transcriptional regulator.
GN VCA0182.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -|- SIMILARITY: CONTAINS A SIGMA-54 FACTOR INTERACTION ATP-BINDING
DOMAIN.
CC EMBL; AE004358; AAF96095.1; .
DR TIGR; VCA0182; .
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003018; GAF.
DR InterPro; IPR002197; HTH_Fis.
DR InterPro; IPR002078; Sig54_interact.
DR Pfam; PF01590; GAF; 1.
DR Pfam; PF00158; Sigma54_activat; 1.
DR PRINTS; PR01590; HTHFIS.
DR QNUMC5; SM00382; AAA; 1.
DR SMART; SM00065; GAF; 1.

```

DR TIGRFAMS: TIGR01199; HTH_fis: 1.
DR PROSITE: PS00675; SIGMA54_INTERACT_1; 1.
DR PROSITE: PS00676; SIGMA54_INTERACT_2; 1.
DR PROSITE: PS00688; SIGMA54_INTERACT_3; 1.
DR PROSITE: PS00045; SIGMA54_INTERACT_4; 1.
DR ATP-binding: DNA-binding; Transcription regulation; Complete proteome:
SQ SEQUENCE 530 AA; 59026 MW; 5B85E12F25A03FA2 CRC64;

Query Match          91.4%; Score 32; DB 16; Length 530;
Best Local Similarity 85.7%; Pred. No. 45;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTNR 7
Db 505 DRANLTNR 511

RESULT 2
Q9D213 PRELIMINARY; PRT; 241 AA.
AC Q9D213;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE Adult male spinal cord cDNA, RIKEN full-length enriched library,
DE clone:A330108A07, full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SPINAL CORD;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Storch K.-F.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK020743; BAB32198.1; -.
DR MGD: P42574; 1PAU.
DR HSSP: P42574; 1PAU.
DR InterPro: IPR002398; ICE.
DR InterPro: IPR002138; ICE_p10.
DR InterPro: IPR001309; ICE_p20.
DR Pfam: PF00655; ICE_p10; 1.
DR Pfam: PF00656; ICE_p20; 1.
DR PRINTS: PR00376; IL1BCENZYM.
DR SMART: SM00115; CASP.1
DR PROSITE: PS01122; CASPASE_CYS; 1.
DR PROSITE: PS01121; CASPASE_HIS; 1.
DR PROSITE: PS02007; CASPASE_P10; 1.
DR PROSITE: PS02008; CASPASE_P20; 1.
SQ SEQUENCE 241 AA; 27834 MW; 53B56B34A3C9609 CRC64;

Query Match          88.6%; Score 31; DB 11; Length 241;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLTNR 7
Db 167 DRKNLTNR 173

RESULT 3
Q9D089 PRELIMINARY; PRT; 276 AA.
AC Q9D089;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

```

```

DE 10 days embryo cDNA, RIKEN full-length enriched library,
DE clone:2610037G10, full insert sequence.
GN CASP6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Storch K.-F.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK011710; BAB27792.1; -.
DR MGD: P42574; 1PAU.
DR HSSP: P42574; 1PAU.
DR InterPro: IPR002398; ICE.
DR InterPro: IPR002138; ICE_p10.
DR InterPro: IPR001309; ICE_p20.
DR Pfam: PF00655; ICE_p10; 1.
DR Pfam: PF00656; ICE_p20; 1.
DR PRINTS: PR00376; IL1BCENZYM.
DR SMART: SM00115; CASP.1
DR PROSITE: PS01122; CASPASE_CYS; 1.
DR PROSITE: PS01121; CASPASE_HIS; 1.
DR PROSITE: PS02007; CASPASE_P10; 1.
DR PROSITE: PS02008; CASPASE_P20; 1.
SQ SEQUENCE 276 AA; 31591 MW; 597EDF2321126B6C CRC64;

Query Match          88.6%; Score 31; DB 11; Length 276;
Best Local Similarity 85.7%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRSNLTNR 7
Db 53 DRDNLTNR 59

RESULT 4
Q99M47 PRELIMINARY; PRT; 276 AA.
AC Q99M47;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Caspase 6.
GN CASP6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; BC002022; AAH02022.1; -.
 DR HSSP; P42574; 1PAU.
 DR MGD; MGI:1312921; Casp6.
 DR InterPro; IPR002398; ICE.
 DR InterPro; IPR002138; ICE_P10.
 DR InterPro; IPR001309; ICE_P20.
 DR Pfam; PF00655; ICE_P10; 1.
 DR Pfam; PF00656; ICE_P20; 1.
 DR PRINTS; PR00376; IL1BCN2YME.
 DR SMART; SM00115; CASG; 1.
 DR PROSITE; PS01122; CASPASE_CYS; 1.
 DR PROSITE; PS01121; CASPASE_HIS; 1.
 DR PROSITE; PS0207; CASPASE_P10; 1.
 DR PROSITE; PS0208; CASPASE_P20; 1.
 SQ SEQUENCE 276 AA; 31563 MW; 5965C5932A127B6C CRC64;

Query Match 88.68; Score 31; DB 11; Length 276;
 Best Local Similarity 85.7%; Pred. No. 40;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRSNLT 7
 || ||||
 Db 53 DRDLNLT 59

RESULT 5
 Q9BQE7

ID Q9BQE7 PRELIMINARY; PRT; 293 AA.
 AC Q9BQE7
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Caspase 6, apoptosis-related cysteine protease.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN 1
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN 2
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC004460; AAH04460.1; -.
 DR EMBL; BC000305; AAH00305.1; -.
 DR HSSP; P42574; 1PAU.
 DR MEROPS; C14.005; -.
 DR InterPro; IPR002398; ICE.
 DR InterPro; IPR002138; ICE_P10.
 DR InterPro; IPR001309; ICE_P20.
 DR Pfam; PF00655; ICE_P10; 1.
 DR Pfam; PF00656; ICE_P20; 1.
 DR PRINTS; PR00376; IL1BCN2YME.
 DR SMART; SM00115; CASG; 1.
 DR PROSITE; PS01122; CASPASE_CYS; 1.
 DR PROSITE; PS01121; CASPASE_HIS; 1.
 DR PROSITE; PS0207; CASPASE_P10; 1.
 DR PROSITE; PS0208; CASPASE_P20; 1.
 KW Protease.
 SQ SEQUENCE 293 AA; 33310 MW; 0738AE4F9791EBD7 CRC64;

Query Match 88.68; Score 31; DB 4; Length 293;
 Best Local Similarity 85.7%; Pred. No. 43;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRSNLT 7
 || ||||
 Db 70 DRDLNLT 76

RESULT 6

P79793
 ID P79793 PRELIMINARY; PRT; 761 AA.
 AC P79793
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Slow myosin heavy chain 2 (Fragment).
 GN SM2.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 OX NCBI_TaxID=9031;
 RN 1
 RP SEQUENCE FROM N.A.
 RC STRAIN=WHITE LEGHORN; TISSUE=LEG MUSCLE;
 RA Chen Q., Moore L.A., Wick M., Bandman E.;
 RT "Identification of a genomic locus containing three slow myosin heavy chain genes";
 RT Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN 2
 RP SEQUENCE FROM N.A.
 RC STRAIN=WHITE LEGHORN; TISSUE=LEG MUSCLE;
 RA Chen Q., Moore L.A., Zhang J., Bumstead N., Bandman E.;
 RT "Fast, slow and ventricular myosin heavy chain genes are found at three different loci";
 RT Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U85023; AAB41890.1; -.
 DR InterPro; IPR002928; Myosin_tail.
 DR InterPro; IPR000533; Tropomyosin.
 DR Pfam; PF01576; Myosin_tail; 1.
 DR PRINTS; PR00194; TROPOMIOSIN.
 DR NON_TER
 FT 1
 SQ SEQUENCE 761 AA; 88289 MW; A8538787E29E3B46 CRC64;

Query Match 88.6%; Score 31; DB 13; Length 761;
 Best Local Similarity 85.7%; Pred. No. 1.e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRSNLT 7
 || ||||
 Db 685 DRKNLT 691

RESULT 7

Q8UWAO
 ID Q8UWAO PRELIMINARY; PRT; 1941 AA.
 AC Q8UWAO
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Myosin heavy chain.
 GN SSMHC.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 OX NCBI_TaxID=9031;
 RN 1
 RP SEQUENCE FROM N.A.
 RA Machida S., Noda S., Takao A., Momma K., Matsuoka R.;
 RT "Differential expression of chick slow skeletal myosin heavy chains in the developing Furkinje myocytes";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB057661; BAB79445.1; -.
 DR InterPro; IPR000048; IQ_region.
 DR InterPro; IPR001609; Myosin_head.
 DR InterPro; IPR004009; Myosin_N.
 DR InterPro; IPR002928; Myosin_tail.
 DR InterPro; IPR000169; SHprot_acsite.
 DR Pfam; PF00612; IQ; 2.

DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS00996; IQ; 1.
DR PROSITE; PS00639; THJOL_PROTEASE_HIS; UNKNOWN_1.
SQ SEQUENCE 1941 AA; 224292 MW; 96B2D8890E148D4D CRC64;

Query Match 88.6%; Score 31; DB 13; Length 1941;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DRSNLTR 7
||| |||||
DB 1865 DRKNLTR 1871

RESULT 8

Q9JX49 PRELIMINARY; PRT; 125 AA.
AC Q9JX49;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Putative lipoprotein.
GN NMA00065.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C., Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T., Davies R.M., Davis P., Davis K., Feltwell T., Hamlin N., Holroyd S., Jagels K., Leather S., Moule S., Mungall K., Quail M.A., Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J., Whitehead S., Barrett B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.";
RL Nature 404:502-506(2000).
DR EMBL; AL162752; CAB83381.1; -.
DR InterPro; IPR002106; AATRNA_ligaseII.
DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
KW Lipoprotein; Complete proteome.
SQ SEQUENCE 125 AA; 13888 MW; A20C7C09F48081E3 CRC64;

Query Match 85.7%; Score 30; DB 16; Length 125;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DRSNLTR 6
||| |||||
DB 88 DRSNLTR 93

RESULT 9

Q9RBT4 PRELIMINARY; PRT; 182 AA.
AC Q9RBT4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE Hypothetical sp. YH102.
OS Pseudomonas 21.0 kDa protein.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Bacteroidetes.
OX NCBI_TaxID=104926;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=YH102;
RA Newman L.M., Zylstra G.J.;
RT "Analysis of genes for p-nitrobenzoate degradation from Pseudomonas sp. strain YH102.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF187880; AAF01449.1; -.
KW Hypothetical protein.
SQ SEQUENCE 182 AA; 20974 MW; B54EB51310BE8E44 CRC64;

Query Match 85.7%; Score 30; DB 2; Length 182;
Best Local Similarity 85.7%; Pred. No. 45;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DRSNLTR 7
||| |||||
DB 34 DRSSLTR 40

RESULT 10

Q9NPY7 PRELIMINARY; PRT; 204 AA.
AC Q9NPY7;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE DJ520B18.2 (FARS1 (Phenylalanine-tRNA synthetase)) (Fragment).
GN DJ520B18.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dunn M.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL121978; CAC00668.1; -.
DR HSSP; P27001; 1PYS.
DR InterPro; IPR002106; AATRNA_ligaseII.
DR InterPro; IPR002319; tRNA-synt_2d.
DR Pfam; PF01409; tRNA-synt_2d; 1.
DR PROSITE; PS00179; AA_TRNA_LIGASE_II_1; UNKNOWN_1.
KW Aminoacyl-tRNA synthetase.
FT NON_TER 204
SQ SEQUENCE 204 AA; 23516 MW; 4DC8AF7BF7FD17B5 CRC64;

Query Match 85.7%; Score 30; DB 4; Length 204;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DRSNLTR 7
||| |||||
DB 62 DHSNLTR 68

RESULT 11

Q96G44 PRELIMINARY; PRT; 292 AA.
AC Q96G44;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Unknown (protein for MGC:16594).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009976; AAH09976.1; -.
DR InterPro; IPR000822; Znf_C2H2.

```
DR Pfam: PF00096; zf-C2H2; 9.
DR ProDom: PD000003; Znf_C2H2; 6.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_8.
DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 9.
KW DNA-binding; Zinc-finger.
SQ SEQUENCE 292 AA; 33497 MW; 3A5E633635C36385 CRC64;

Query Match 85.7%; Score 30; DB 4; Length 292;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRSNLTR 7
Db 135 DRSLSLR 141

RESULT 12
Q9BSG1 PRELIMINARY; PRT; 292 AA.
AC Q9BSG1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created).
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 33.5 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: BC005068; AAH05068.1; -.
DR HSSP: P07248; 2ADR.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF00096; zf-C2H2; 9.
DR PRINTS: PR00048; ZINC_FINGER.
DR ProDom: PD000003; Znf_C2H2; 6.
DR SMART: SM00355; Znf_C2H2; 9.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 8.
DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 9.
KW DNA-binding; Hypothetical protein; Metal-binding; Nuclear protein;
KW Zinc-finger.
SQ SEQUENCE 292 AA; 33471 MW; 2B4F632724C36385 CRC64;

Query Match 85.7%; Score 30; DB 4; Length 292;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRSNLTR 7
Db 135 DRSLSLR 141

RESULT 13
Q52190 PRELIMINARY; PRT; 321 AA.
ID Q52190;
AC Q52190;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE TRAA.
GN TRAA.
OS Enterococcus faecalis (Streptococcus faecalis).
OG plasmid pPDI.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96032393; PubMed=7559344;
RA Nakayama J., Yoshida K., Kobayashi H., Isogai A., Clewell D.,
RT "Cloning and characterization of a region of enterococcus faecalis
RT plasmid pPDI encoding pheromone inhibitor (ipd), pheromone sensitivity
RT J. Bacteriol. 177:5567-5573(1995).
RL J. Bacteriol. 177:5574-5581(1995).
DR EMBL: D78016; BAA11197.1; -.
KW Plasmid.
SQ SEQUENCE 321 AA; 37700 MW; 921B9D176B72B2C4 CRC64;

Query Match 85.7%; Score 30; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRSNLTR 6
Db 52 DRSNLTR 57

RESULT 14
Q52195 PRELIMINARY; PRT; 321 AA.
ID Q52195;
AC Q52195;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE TRAA.
GN TRAA.
OS Enterococcus faecalis (Streptococcus faecalis).
OG plasmid pPDI.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96032393; PubMed=7559344;
RA Nakayama J., Yoshida K., Kobayashi H., Isogai A., Clewell D.,
RT "Cloning and characterization of a region of enterococcus faecalis
RT plasmid pPDI encoding pheromone inhibitor (ipd), pheromone sensitivity
RT (trac), and pheromone shutdown (trab) genes."
RL J. Bacteriol. 177:5567-5573(1995).
DR EMBL: D78016; BAA11197.1; -.
KW Plasmid.
SQ SEQUENCE 321 AA; 37700 MW; 921B9D176B72B2C4 CRC64;

Query Match 85.7%; Score 30; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRSNLTR 6
Db 52 DRSNLTR 57

RESULT 15
Q9UMC5 PRELIMINARY; PRT; 395 AA.
ID Q9UMC5;
AC Q9UMC5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Zinc finger 2.2 (Fragment).
GN ZNF2.2.
OS Homo sapiens (Human).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92051312; PubMed=1945843;
 RA Rosati M., Marino M., Franze A., Tramontano A., Grimaldi G.;
 RT "Members of the zinc finger protein gene family sharing a conserved N-
 RT Terminal module";
 RL Nucleic Acids Res. 19:5661-5667(1991).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; X60152; CAB52138.1; -;
 DR HSSP; P07248; 2ADR.
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR000822; znf_C2H2.
 DR Pfam; PF01352; KRAB; 1.
 DR Pfam; PF00096; zf-C2H2; 8.
 DR PRINTS; PR00048; ZINC_FINGER.
 DR ProDom; PD000003; znf_C2H2; 6.
 DR SMART; SM00349; KRAB; 1.
 DR SMART; SM00355; znf_C2H2; 8.
 DR PROSITE; PS08005; KRAB; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 8.
 KW DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
 FT NON_TER 1
 FT NON_TER 395
 SQ SEQUENCE 395 AA; 45324 MW; 63FBD3DF605A1393 CRC64;

Query Match 85.7%; Score 30; DB 4; Length 395;
 Best Local Similarity 85.7%; Pred. No. 99;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLR 7
 Db 271 DRSSLTR 277

Search completed: July 15, 2003, 12:54:47
 Job time : 25.6667 secs

GenCore version 5.1.6
Copyright.(c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:53:06 ; Search time 16.3333 Seconds
(without alignments)
49.894 Million cell updates/sec

Title: US-10-006-069A-55

Perfect score: 35

Sequence: 1 DRSNLTR 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	35	100.0	7	9	US-10-006-069A-36
2	35	100.0	7	9	Sequence 36, Appl
3	35	100.0	7	9	Sequence 55, Appl
4	35	100.0	7	9	Sequence 56, Appl
5	35	100.0	7	9	Sequence 59, Appl
6	35	100.0	7	9	Sequence 67, Appl
7	35	100.0	7	9	US-10-006-069A-134
8	35	100.0	7	9	Sequence 134, App
9	35	100.0	7	9	Sequence 167, App
10	35	100.0	7	9	Sequence 170, App
11	35	100.0	7	9	Sequence 196, App
12	35	100.0	7	9	Sequence 198, App
13	35	100.0	7	9	Sequence 36, Appl
14	35	100.0	7	9	Sequence 55, Appl
15	35	100.0	7	9	Sequence 56, Appl
16	35	100.0	7	9	Sequence 59, Appl
17	35	100.0	7	9	Sequence 67, Appl
18	35	100.0	7	9	Sequence 134, App
19	35	100.0	7	9	Sequence 167, App
					Sequence 170, App
					Sequence 196, App

Sequence 198, App
Sequence 37, Appl
Sequence 395, App
Sequence 1376, Ap
Sequence 1454, Ap
Sequence 1464, Ap
Sequence 1525, Ap
Sequence 1553, Ap
Sequence 1559, Ap
Sequence 1564, Ap
Sequence 1577, Ap
Sequence 1899, Ap
Sequence 2668, Ap
Sequence 2705, Ap
Sequence 2737, Ap
Sequence 2791, Ap
Sequence 2792, Ap
Sequence 2887, Ap
Sequence 2995, Ap
Sequence 2996, Ap
Sequence 2999, Ap
Sequence 3006, Ap
Sequence 3131, Ap
Sequence 3197, Ap
Sequence 3216, Ap
Sequence 3286, Ap

ALIGNMENTS

RESULT 1

US-10-006-069A-36

; Sequence 36, Application US/10006069A

; Publication No. US20030021776A1

; GENERAL INFORMATION:

; APPLICANT: Rebar, Edward

; APPLICANT: Jamieson, Andrew

; APPLICANT: Liu, Qiang

; APPLICANT: Liu, Pei-Qi

; APPLICANT: Wolffe, Alan

; APPLICANT: Eisenberg, Stephen P.

; APPLICANT: Jarvis, Eric

; APPLICANT: Sangamo BioSciences, Inc.

; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc

; FILE REFERENCE: 019496-00583005

; CURRENT APPLICATION NUMBER: US/10/006,069A

; CURRENT FILING DATE: 2001-12-17

; PRIOR APPLICATION NUMBER: US 09/733,604

; PRIOR FILING DATE: 2000-12-07

; PRIOR APPLICATION NUMBER: US 09/736,083

; PRIOR FILING DATE: 2000-12-12

; PRIOR APPLICATION NUMBER: US 09/846,033

; PRIOR FILING DATE: 2001-04-30

; NUMBER OF SEQ ID NOS: 252

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 36

; LENGTH: 7

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: finger

US-10-006-069A-36

Query Match 100.0%; Score 35; DB 9; Length 7;

Best Local Similarity 100.0%; Pred. No. 4e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRSNLTR 7

Db 1 DRSNLTR 7

RESULT 2

US-10-006-069A-55
; Sequence 55, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006,069A
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 55
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-55

Query Match 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
Db 1 DRSNLTR 7

RESULT 3

US-10-006-069A-56
; Sequence 56, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006,069A
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 56
; LENGTH: 7
; TYPE: PRT

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-56

Query Match. 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
Db 1 DRSNLTR 7

RESULT 4

US-10-006-069A-59
; Sequence 59, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006,069A
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-59

Query Match 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
Db 1 DRSNLTR 7

RESULT 5

US-10-006-069A-67
; Sequence 67, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US

; CURRENT APPLICATION NUMBER: US/10/006.069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2000-12-07 US 09/733,604
; PRIOR FILING DATE: 2000-12-07 US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR FILING DATE: 2000-12-12 US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 67
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-67

Query Match 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 7
Db 1 DRSNLT 7

RESULT 6
US-10-006-069A-134
; Sequence 134, Application US/10006069A
; Publication No. US20030021776A1

; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006.069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2000-12-07 US 09/733,604
; PRIOR FILING DATE: 2000-12-07 US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 134
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-134

Query Match 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 7
Db 1 DRSNLT 7

RESULT 7
US-10-006-069A-167
; Sequence 167, Application US/10006069A

; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006.069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2000-12-07 US 09/733,604
; PRIOR FILING DATE: 2000-12-07 US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 167
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-167

Query Match 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLT 7
Db 1 DRSNLT 7

RESULT 8
US-10-006-069A-170
; Sequence 170, Application US/10006069A
; Publication No. US20030021776A1

; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006.069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2000-12-07 US 09/733,604
; PRIOR FILING DATE: 2000-12-07 US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 170
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-170

Query Match 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
| | | | | | |
DB 1 DRSNLTR 7

RESULT 9
US-10-006-069A-196
; Sequence 196, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolfe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006,069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 196
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: recognition helix
US-10-006-069A-196

Query Match 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
| | | | | | |
DB 1 DRSNLTR 7

RESULT 10
US-10-006-069A-198
; Sequence 198, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolfe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006,069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2000-12-07

; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 198
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: recognition helix
US-10-006-069A-198

Query Match 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
| | | | | | |
DB 1 DRSNLTR 7

RESULT 11
US-09-846-033B-36
; Sequence 36, Application US/09846033B
; Publication No. US20030044404A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolfe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005820US
; CURRENT APPLICATION NUMBER: US/09/846,033B
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-09-846-033B-36

Query Match 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
| | | | | | |
DB 1 DRSNLTR 7

RESULT 12
US-09-846-033B-55
; Sequence 55, Application US/09846033B
; Publication No. US20030044404A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi

; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005820US
; CURRENT APPLICATION NUMBER: US/09/846,033B
; CURRENT FILING DATE: 2001-04-30
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 55
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-09-846-033B-55

Query Match 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. NO. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
Db 1 DRSNLTR 7

RESULT 13
US-09-846-033B-56
; Sequence 56, Application US/09846033B
; Publication No. US2003004404A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005820US
; CURRENT APPLICATION NUMBER: US/09/846,033B
; CURRENT FILING DATE: 2001-04-30
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 56
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-09-846-033B-56

Query Match 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. NO. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
Db 1 DRSNLTR 7

RESULT 14
US-09-846-033B-59
; Sequence 59, Application US/09846033B
; Publication No. US2003004404A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005820US
; CURRENT APPLICATION NUMBER: US/09/846,033B
; CURRENT FILING DATE: 2001-04-30
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-09-846-033B-59

Query Match 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. NO. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
Db 1 DRSNLTR 7

RESULT 15
US-09-846-033B-67
; Sequence 67, Application US/09846033B
; Publication No. US2003004404A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005820US
; CURRENT APPLICATION NUMBER: US/09/846,033B
; CURRENT FILING DATE: 2001-04-30
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 67
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-09-846-033B-67

Query Match 100.0%; Score 35; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRSNLTR 7
Db 1 DRSNLTR 7

Search completed: July 15, 2003, 12:57:03
Job time : 17.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:53:06 ; Search time 16.3333 Seconds
(without alignments)
49.894 Million cell updates/sec

Title: US-10-006-069A-68
Perfect score: 36
Sequence: 1 RSDHLR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 11641973 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	7	US-10-006-069A-64	Sequence 64, Appl
2	36	100.0	7	US-10-006-069A-68	Sequence 68, Appl
3	36	100.0	7	US-10-006-069A-91	Sequence 91, Appl
4	36	100.0	7	US-10-006-069A-101	Sequence 101, Appl
5	36	100.0	7	US-10-006-069A-102	Sequence 102, Appl
6	36	100.0	7	US-10-006-069A-103	Sequence 103, Appl
7	36	100.0	7	US-10-006-069A-104	Sequence 104, Appl
8	36	100.0	7	US-10-006-069A-105	Sequence 105, Appl
9	36	100.0	7	US-10-006-069A-106	Sequence 106, Appl
10	36	100.0	7	US-10-006-069A-109	Sequence 109, Appl
11	36	100.0	7	US-10-006-069A-111	Sequence 111, Appl
12	36	100.0	7	US-10-006-069A-113	Sequence 113, Appl
13	36	100.0	7	US-10-006-069A-114	Sequence 114, Appl
14	36	100.0	7	US-10-006-069A-116	Sequence 116, Appl
15	36	100.0	7	US-10-006-069A-154	Sequence 154, Appl
16	36	100.0	7	US-10-006-069A-163	Sequence 163, Appl
17	36	100.0	7	US-10-006-069A-169	Sequence 169, Appl
18	36	100.0	7	US-10-006-069A-172	Sequence 172, Appl
19	36	100.0	7	US-10-006-069A-174	Sequence 174, Appl

20	36	100.0	7	US-10-006-069A-177	Sequence 177, App
21	36	100.0	7	US-10-006-069A-200	Sequence 200, App
22	36	100.0	7	US-09-846-033B-64	Sequence 64, Appl
23	36	100.0	7	US-09-846-033B-68	Sequence 68, Appl
24	36	100.0	7	US-09-846-033B-91	Sequence 91, Appl
25	36	100.0	7	US-09-846-033B-101	Sequence 101, App
26	36	100.0	7	US-09-846-033B-102	Sequence 102, App
27	36	100.0	7	US-09-846-033B-103	Sequence 103, App
28	36	100.0	7	US-09-846-033B-104	Sequence 104, App
29	36	100.0	7	US-09-846-033B-105	Sequence 105, App
30	36	100.0	7	US-09-846-033B-106	Sequence 106, App
31	36	100.0	7	US-09-846-033B-109	Sequence 109, App
32	36	100.0	7	US-09-846-033B-111	Sequence 111, App
33	36	100.0	7	US-09-846-033B-113	Sequence 113, App
34	36	100.0	7	US-09-846-033B-114	Sequence 114, App
35	36	100.0	7	US-09-846-033B-116	Sequence 116, App
36	36	100.0	7	US-09-846-033B-154	Sequence 154, App
37	36	100.0	7	US-09-846-033B-163	Sequence 163, App
38	36	100.0	7	US-09-846-033B-169	Sequence 169, App
39	36	100.0	7	US-09-846-033B-172	Sequence 172, App
40	36	100.0	7	US-09-846-033B-174	Sequence 174, App
41	36	100.0	7	US-09-846-033B-177	Sequence 177, App
42	36	100.0	7	US-09-846-033B-200	Sequence 200, App
43	36	100.0	7	US-10-055-713-62	Sequence 62, Appl
44	36	100.0	7	US-10-084-826-5	Sequence 5, Appl
45	36	100.0	7	US-09-990-186-229	Sequence 229, App

ALIGNMENTS

RESULT 1
US-10-006-069A-64
; Sequence 64, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-00583003
; CURRENT APPLICATION NUMBER: US/10/006,069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-64

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RSDHLR 7

Db 1 RSDHLR 7

RESULT 2
US-10-006-069A-68
; Sequence 68, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE OF INVENTION: Finger Proteins
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006, 069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 68
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-68

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 1 RSDHLSR 7

RESULT 3
US-10-006-069A-91
; Sequence 91, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE OF INVENTION: Finger Proteins
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006, 069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 7
; TYPE: PRT

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-91

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 1 RSDHLSR 7

RESULT 4
US-10-006-069A-101
; Sequence 101, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE OF INVENTION: Finger Proteins
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006, 069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 101
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-101

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
Db 1 RSDHLSR 7

RESULT 5
US-10-006-069A-102
; Sequence 102, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE OF INVENTION: Finger Proteins
; FILE REFERENCE: 019496-005830US

; CURRENT APPLICATION NUMBER: US/10/006.069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07 US 09/736,083
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 102
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-102

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSDHLSR 7
Db 1 RSDHLSR 7

RESULT 6
US-10-006-069A-103
; Sequence 103, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolfe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006.069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07 US 09/736,083
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 103
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-103

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSDHLSR 7
Db 1 RSDHLSR 7

RESULT 7
US-10-006-069A-104
; Sequence 104, Application US/10006069A

; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolfe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006.069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07 US 09/736,083
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-104

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSDHLSR 7
Db 1 RSDHLSR 7

RESULT 8
US-10-006-069A-105
; Sequence 105, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolfe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006.069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07 US 09/736,083
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 105
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-105

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
 |||||||
Db 1 RSDHLSR 7

RESULT 9

US-10-006-069A-106
; Sequence 106, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006,069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 106
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-106

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
 |||||||
Db 1 RSDHLSR 7

RESULT 10

US-10-006-069A-109
; Sequence 109, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006,069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07

; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 109
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-109

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
 |||||||
Db 1 RSDHLSR 7

RESULT 11

US-10-006-069A-111
; Sequence 111, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006,069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 111
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-111

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
 |||||||
Db 1 RSDHLSR 7

RESULT 12

US-10-006-069A-113
; Sequence 113, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew

APPLICANT: Liu, Qiang
APPLICANT: Liu, Pei-Qi
APPLICANT: Wolffe, Alan
APPLICANT: Eisenberg, Stephen P.
APPLICANT: Jarvis, Eric
APPLICANT: Sangamo BioSciences, Inc.
TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
FILE REFERENCE: 019496-005830US
CURRENT APPLICATION NUMBER: US/10/006,069A
CURRENT FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: US 09/733,604
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 09/736,083
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: US 09/846,033
PRIOR FILING DATE: 2001-04-30
NUMBER OF SEQ ID NOS: 252
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 113
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: finger
US-10-006-069A-113

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
|||||||
Db 1 RSDHLSR 7

RESULT 13
US-10-006-069A-114
Sequence 114, Application US/10006069A
Publication No. US20030021776A1
GENERAL INFORMATION:
APPLICANT: Rebar, Edward
APPLICANT: Jamieson, Andrew
APPLICANT: Liu, Qiang
APPLICANT: Liu, Pei-Qi
APPLICANT: Wolffe, Alan
APPLICANT: Eisenberg, Stephen P.
APPLICANT: Jarvis, Eric
APPLICANT: Sangamo BioSciences, Inc.
TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
FILE REFERENCE: 019496-005830US
CURRENT APPLICATION NUMBER: US/10/006,069A
CURRENT FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: US 09/733,604
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 09/736,083
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: US 09/846,033
PRIOR FILING DATE: 2001-04-30
NUMBER OF SEQ ID NOS: 252
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 114
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: finger
US-10-006-069A-114

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
|||||||
Db 1 RSDHLSR 7

RESULT 14
US-10-006-069A-116
Sequence 116, Application US/10006069A
Publication No. US20030021776A1
GENERAL INFORMATION:
APPLICANT: Rebar, Edward
APPLICANT: Jamieson, Andrew
APPLICANT: Liu, Qiang
APPLICANT: Liu, Pei-Qi
APPLICANT: Wolffe, Alan
APPLICANT: Eisenberg, Stephen P.
APPLICANT: Jarvis, Eric
APPLICANT: Sangamo BioSciences, Inc.
TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
FILE REFERENCE: 019496-005830US
CURRENT APPLICATION NUMBER: US/10/006,069A
CURRENT FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: US 09/733,604
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 09/736,083
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: US 09/846,033
PRIOR FILING DATE: 2001-04-30
NUMBER OF SEQ ID NOS: 252
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 116
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: finger
US-10-006-069A-116

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
|||||||
Db 1 RSDHLSR 7

RESULT 15
US-10-006-069A-154
Sequence 154, Application US/10006069A
Publication No. US20030021776A1
GENERAL INFORMATION:
APPLICANT: Rebar, Edward
APPLICANT: Jamieson, Andrew
APPLICANT: Liu, Qiang
APPLICANT: Liu, Pei-Qi
APPLICANT: Wolffe, Alan
APPLICANT: Eisenberg, Stephen P.
APPLICANT: Jarvis, Eric
APPLICANT: Sangamo BioSciences, Inc.
TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
FILE REFERENCE: 019496-005830US
CURRENT APPLICATION NUMBER: US/10/006,069A
CURRENT FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: US 09/733,604
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 09/736,083
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: US 09/846,033
PRIOR FILING DATE: 2001-04-30

; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 154
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-154

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSDHLSR 7
| | | | | | |
Db 1 RSDHLSR 7

Search completed: July 15, 2003, 12:57:03
Job time : 16.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:48:06 ; Search time 33.3333 Seconds
(without alignments)
27.983 Million cell updates/sec

Title: US-10-006-069A-247
Perfect score: 36
Sequence: 1 TSGHLSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	7	ABJ03962	Human VEGF-targete
2	36	100.0	7	ABP48959	Zinc finger protei
3	33	91.7	7	ABJ03960	Human VEGF-targete
4	33	91.7	7	ABJ03963	Human VEGF-targete
5	33	91.7	7	ABP48380	Zinc finger protei
6	33	91.7	7	ABP48797	Zinc finger protei
7	33	91.7	7	ABP48798	Zinc finger protei
8	33	91.7	7	ABP48802	Zinc finger protei
9	33	91.7	7	ABP49078	Zinc finger protei
10	33	91.7	7	ABP49085	Zinc finger protei

11	33	91.7	7	23	ABP49088	Zinc finger protei
12	33	91.7	7	23	ABP49102	Zinc finger protei
13	32	88.9	7	23	ABP48958	Zinc finger protei
14	32	88.9	284	21	AAG14909	Arabidopsis thalia
15	32	88.9	285	21	AAG14907	Arabidopsis thalia
16	32	88.9	302	21	AAG14908	Arabidopsis thalia
17	32	88.9	302	21	AAG14908	Arabidopsis thalia
18	32	88.9	303	21	AAG2543	A. thaliana transc
19	32	88.9	302	21	AAG45426	Arabidopsis thalia
20	32	88.9	356	22	AAE00655	Human protein tyro
21	32	88.9	542	23	AAU79279	Human serine-threo
22	32	88.9	565	22	AAE65692	Novel protein kina
23	32	88.9	580	23	AAU79280	Human serine-threo
24	31	86.1	7	23	ABJ03964	Human VEGF-targete
25	31	86.1	7	23	ABP48743	Zinc finger protei
26	31	86.1	7	23	ABP48746	Zinc finger protei
27	31	86.1	7	23	ABP48749	Zinc finger protei
28	31	86.1	7	23	ABP48752	Zinc finger protei
29	31	86.1	7	23	ABP48755	Zinc finger protei
30	31	86.1	7	23	ABP48758	Zinc finger protei
31	31	86.1	7	23	ABP48761	Zinc finger protei
32	31	86.1	7	23	ABP48764	Zinc finger protei
33	31	86.1	7	23	ABP49007	Zinc finger protei
34	31	86.1	7	23	ABP49010	Zinc finger protei
35	31	86.1	7	23	ABP49019	Zinc finger protei
36	31	86.1	52	22	AAU46550	Propionibacterium
37	31	86.1	137	22	AAU67542	Propionibacterium
38	31	86.1	173	21	AAE23358	Arabidopsis thalia
39	31	86.1	173	21	AAE23357	Arabidopsis thalia
40	31	86.1	196	21	AAE23356	Arabidopsis thalia
41	31	86.1	199	21	AAE23356	Arabidopsis thalia
42	31	86.1	229	21	AAE22824	Arabidopsis thalia
43	31	86.1	229	21	AAE22824	Arabidopsis thalia
44	31	86.1	229	21	AAE22824	Arabidopsis thalia
45	31	86.1	254	21	AAE22823	Arabidopsis thalia

ALIGNMENTS

RESULT 1

ABJ03962
ID ASJ03962 standard; Peptide; 7 AA.

XX AC ASJ03962;

XX DT 25-SEP-2002 (first entry)

XX DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 247.

XX KW Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
XX KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
XX KW gene therapy; antidiabetic; antidiabetic; antidiabetic; antidiabetic;
XX KW antidiabetic; antidiabetic; antidiabetic; antidiabetic;
XX KW osteopathic; antidiabetic; antidiabetic; antidiabetic;
XX OS Homo sapiens.

XX PN WC200246412-A2.

XX PD 13-JUN-2002.

XX PF 06-DEC-2001; 2001WO-US46861.

XX PR 07-DEC-2000; 2000US-0733604.

XX PR 12-DEC-2000; 2000US-0736083.

XX PR 30-APR-2001; 2001US-0846033.

XX PA (SANG-) SANGAMO BIOSCIENCES INC.

XX PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;

XX PI Jarvis E;

XX XX

DR WPI; 2002-527918/56.
 XX New zinc finger protein that binds to target site in vascular
 PT endothelial growth factor gene, useful for modulating expression of the
 PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
 PT ulcer ..
 XX
 PS Claim 4; Page 103; 195pp; English.
 XX
 XX The present invention relates to a zinc finger protein that binds to a
 CC target site in one or more vascular endothelial growth factor (VEGF)
 CC genes. The protein is useful for modulating expression of a VEGF gene,
 CC thereby regulating angiogenesis and vasculogenesis. This can be used to
 CC treat atherosclerosis, ischemia, arthritis, wounds, ulcers, tumours,
 CC diabetic retinopathy or psoriasis. The present sequence is a peptide
 CC shown in the invention.
 XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 36; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TSGHLSR 7
 DB 1 TSGHLSR 7
 RESULT 2
 ABP48959
 ID ABP48959 standard; Peptide; 7 AA.
 XX
 XX ABP48959;
 XX
 XX 28-AUG-2002 (first entry)
 DT
 XX
 DE Zinc finger protein related peptide motif SEQ ID NO:1201.
 XX
 KW Zinc finger protein; ZFP; DNA binding protein; zinc finger.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200242459-A2.
 XX
 XX 30-MAY-2002.
 PD
 XX
 XX 20-NOV-2001; 2001WO-US43438.
 PF
 XX
 XX 20-NOV-2000; 2000US-0716637.
 PR
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 PA
 XX
 PI Liu Q;
 XX
 DR WPI; 2002-500284/53.
 XX
 XX New zinc finger protein that binds to target site, useful in studying
 PT gene function and for human therapeutics and plant engineering,
 PT comprises first, second and third zinc fingers, ordered from N- to
 PT C-terminus ..
 XX
 PS Claim 1; Page 44; 81pp; English.
 XX
 XX The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target subsite. Also described are: (1) a polypeptide
 CC (I1) comprising (I); (2) a polynucleotide (II) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target subsite, selecting the F2 zinc finger such
 CC that it binds to the S2 target subsite, and selecting the F3 zinc

CC finger such that it binds to the S3 target subsite, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target subsites having the nucleotide G in the 5'-most position of the
 CC subsite. (I) is useful in studying gene function, and for human
 CC therapeutic and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determine the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.
 XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 36; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TSGHLSR 7
 DB 1 TSGHLSR 7
 RESULT 3
 ABJ03960
 ID ABJ03960 standard; Peptide; 7 AA.
 XX
 XX ABJ03960;
 AC
 XX
 DT 25-SEP-2002 (first entry)
 XX
 DE Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 245.
 XX
 KW Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
 KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
 KW gene therapy; antiatherosclerotic; vasotrophic; antiarthritic; vulnary;
 KW antitumor; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
 KW osteopathic; antiinfertility.
 XX
 OS Homo sapiens.
 XX
 PN WO200246412-A2.
 XX
 XX 13-JUN-2002.
 PD
 XX
 XX 06-DEC-2001; 2001WO-US46861.
 PF
 XX
 XX 07-DEC-2000; 2000US-0733604.
 PR
 XX 12-DEC-2000; 2000US-0736083.
 PR
 XX 30-APR-2001; 2001US-0846033.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 XX Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
 PI Jarvis E;
 XX
 DR WPI; 2002-527918/56.
 XX
 XX New zinc finger protein that binds to target site in vascular
 PT endothelial growth factor gene, useful for modulating expression of the
 PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
 PT ulcer ..
 XX
 PS Claim 4; Page 103; 195pp; English.
 XX
 XX The present invention relates to a zinc finger protein that binds to a
 CC target site in one or more vascular endothelial growth factor (VEGF)
 CC genes. The protein is useful for modulating expression of a VEGF gene,
 CC thereby regulating angiogenesis and vasculogenesis. This can be used to
 CC treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
 CC diabetic retinopathy or psoriasis. The present sequence is a peptide

CC shown in the invention.

XX Sequence 7 AA;

Query Match 91.7%; Score 33; DB 23; Length 7;
Best Local Similarity 85.7%; Pred. No. 7.8e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TSGHLSR 7
|||||:
Db 1 TSGHLTR 7

RESULT 4

ABJ03963
ID ABJ03963 standard; Peptide; 7 AA.

XX AC
XX ABJ03963;

XX 25-SEP-2002 (first entry)

XX Human VEGF-targeted zinc finger protein fragment SEQ ID NO: 248.

XX zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
XX diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
XX gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnary;
XX antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
XX osteopathic; antinfertility.

OS Homo sapiens.

XX WO200246412-A2.

XX 13-JUN-2002.

XX 06-DEC-2001; 2001WO-US46861.

XX 07-DEC-2000; 2000US-0733604.

XX 12-DEC-2000; 2000US-0736083.

XX 30-APR-2001; 2001US-0846033.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Rebar E, Jamieson A, Liu Q, Wolfe A, Eisenberg SP;

XX Jarvis E;

XX WPI; 2002-527918/56.

XX New zinc finger protein that binds to target site in vascular

XX endothelial growth factor gene, useful for modulating expression of the
XX gene and for treating atherosclerosis, ischemia, arthritis, wound or
XX ulcer

XX Claim 4; Page 103; 195pp; English.

XX The present invention relates to a zinc finger protein that binds to a
XX target site in one or more vascular endothelial growth factor (VEGF)
XX genes. The protein is useful for modulating expression of a VEGF gene,
XX thereby regulating angiogenesis and vasculogenesis. This can be used to
XX treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
XX diabetic retinopathy or psoriasis. The present sequence is a peptide
XX shown in the invention.

XX Sequence 7 AA;

Query Match 91.7%; Score 33; DB 23; Length 7;
Best Local Similarity 85.7%; Pred. No. 7.8e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TSGHLSR 7
|||||:
Db 1 TSGHLAR 7

RESULT 5

ABP48380
ID ABP48380 standard; Peptide; 7 AA.

XX AC
XX ABP48380;

XX 28-AUG-2002 (first entry)

XX Zinc finger protein related peptide motif SEQ ID NO:393.

XX Zinc finger protein; ZFP; DNA binding protein; zinc finger.

OS Homo sapiens.

XX Synthetic.

XX WO200242459-A2.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US43438.

XX 20-NOV-2000; 2000US-0716637.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Liu Q;

XX WPI; 2002-500284/53.

XX New zinc finger protein that binds to target site, useful in studying
XX gene function and for human therapeutics and plant engineering, to
XX comprises first, second and third zinc fingers, ordered from N-
XX C-terminus

XX Example 1; Page 37; 81pp; English.

XX The present invention describes a zinc finger protein (I) that binds to
XX a target site, comprising a first (F1), a second (F2), and a third (F3)
XX zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
XX target site comprises, in 3'-5' direction, a first (S1), a second (S2),
XX and a third (S3) target subsite. Also described are: (I) a polypeptide
XX (I) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
XX (3) designing (M) (I) involves selecting the F1 zinc finger such that
XX it binds to the S1 target subsite, selecting the F2 zinc finger such
XX that it binds to the S2 target subsite, and selecting the F3 zinc
XX finger such that it binds to the S3 target subsite, thus designing (I)
XX that binds to a target site. (I) is useful for recognition of triplet
XX target subsites having the nucleotide G in the 5'-most position of the
XX subsite. (I) is useful in studying gene function, and for human
XX therapeutics and plant engineering. (I), (II) or (III) is useful in
XX therapeutic methods to modulate the expression of a target region within
XX a subject, in diagnostic methods for sequence specific detection of
XX target nucleic acid in a sample, and in assays to determine the
XX phenotype and function of gene expression. (I) has improved affinity
XX and specificity for their target sequences, as well as enhanced
XX biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
XX represent DNA target sequences and zinc finger peptides which are given
XX in the exemplification of the present invention.

XX Sequence 7 AA;

Query Match 91.7%; Score 33; DB 23; Length 7;
Best Local Similarity 85.7%; Pred. No. 7.8e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TSGHLSR 7
|||||:
Db 1 TSGHLTR 7

RESULT 6

ABP48797

ID XX ABP48797 standard; Peptide; 7 AA.
 AC XX ABP48797;
 DT XX
 DE XX 28-AUG-2002 (first entry)
 XX XX Zinc finger protein related peptide motif SEQ ID NO:1147.
 XX XX Zinc finger protein; ZFP; DNA binding protein; zinc finger.
 XX XX Homo sapiens.
 OS XX Synthetic.
 OS XX WO200242459-A2.
 PN XX 30-MAY-2002.
 PD XX
 XX XX 20-NOV-2001; 2001WO-US43438.
 XX XX 20-NOV-2000; 2000US-0716637.
 XX XX (SANG-) SANGAMO BIOSCIENCES INC.
 PA XX
 PI XX Liu Q;
 DR XX WPI; 2002-500284/53.
 XX XX
 XX XX New zinc finger protein that binds to target site, useful in studying
 XX PT gene function and for human therapeutics and plant engineering,
 XX PT comprises first, second and third zinc fingers, ordered from N- to
 XX PT C-terminus
 XX PS Example 1; Page 42; 81pp; English.
 XX XX
 XX CC The present invention describes a zinc finger protein (I) that binds to
 XX CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 XX CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 XX CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 XX CC and a third (S3) target subsite. Also described are: (1) a polypeptide
 XX CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 XX CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 XX CC it binds to the S1 target subsite, selecting the F2 zinc finger such
 XX CC that it binds to the S2 target subsite, and selecting the F3 zinc
 XX CC finger such that it binds to the S3 target subsite, thus designing (I)
 XX CC that binds to a target site. (I) is useful for recognition of triplet
 XX CC target subsites having the nucleotide G in the 5'-most position of the
 XX CC subsite. (I) is useful in studying gene function, and for human
 XX CC therapeutic and plant engineering. (I), (II) or (III) is useful in
 XX CC therapeutic methods to modulate the expression of a target region within
 XX CC a subject, in diagnostic methods for sequence specific detection of
 XX CC target nucleic acid in a sample, and in assays to determine the
 XX CC phenotype and function of gene expression. (I) has improved affinity
 XX CC and specificity for their target sequences, as well as enhanced
 XX CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
 XX CC represent DNA target sequences and zinc finger peptides which are given
 XX CC in the exemplification of the present invention.
 XX SQ Sequence 7 AA;
 Query Match 91.7%; Score 33; DB 23; Length 7;
 Best Local Similarity 85.7%; Pred. No. 7.8e+05;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TSGHLSR 7
 Db 1 TSGHLTR 7
 RESULT 7
 ABP48798
 ID ABP48798 standard; Peptide; 7 AA.
 XX XX
 AC XX ABP48798;

XX 28-AUG-2002 (first entry)
 DT XX
 DE XX Zinc finger protein related peptide motif SEQ ID NO:738.
 XX XX Zinc finger protein; ZFP; DNA binding protein; zinc finger.
 KW XX Homo sapiens.
 OS XX Synthetic.
 OS XX WO200242459-A2.
 PN XX 30-MAY-2002.
 PD XX
 XX XX 20-NOV-2001; 2001WO-US43438.
 XX XX 20-NOV-2000; 2000US-0716637.
 XX XX (SANG-) SANGAMO BIOSCIENCES INC.
 PA XX
 PI XX Liu Q;
 DR XX WPI; 2002-500284/53.
 XX XX
 XX XX New zinc finger protein that binds to target site, useful in studying
 XX PT gene function and for human therapeutics and plant engineering,
 XX PT comprises first, second and third zinc fingers, ordered from N- to
 XX PT C-terminus
 XX PS Example 1; Page 42; 81pp; English.
 XX XX
 XX CC The present invention describes a zinc finger protein (I) that binds to
 XX CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 XX CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 XX CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 XX CC and a third (S3) target subsite. Also described are: (1) a polypeptide
 XX CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 XX CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 XX CC it binds to the S1 target subsite, selecting the F2 zinc finger such
 XX CC that it binds to the S2 target subsite, and selecting the F3 zinc
 XX CC finger such that it binds to the S3 target subsite, thus designing (I)
 XX CC that binds to a target site. (I) is useful for recognition of triplet
 XX CC target subsites having the nucleotide G in the 5'-most position of the
 XX CC subsite. (I) is useful in studying gene function, and for human
 XX CC therapeutic and plant engineering. (I), (II) or (III) is useful in
 XX CC therapeutic methods to modulate the expression of a target region within
 XX CC a subject, in diagnostic methods for sequence specific detection of
 XX CC target nucleic acid in a sample, and in assays to determine the
 XX CC phenotype and function of gene expression. (I) has improved affinity
 XX CC and specificity for their target sequences, as well as enhanced
 XX CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
 XX CC represent DNA target sequences and zinc finger peptides which are given
 XX CC in the exemplification of the present invention.
 XX SQ Sequence 7 AA;
 Query Match 91.7%; Score 33; DB 23; Length 7;
 Best Local Similarity 85.7%; Pred. No. 7.8e+05;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TSGHLSR 7
 Db 1 TSGHLTR 7
 RESULT 8
 ABP48802
 ID ABP48802 standard; Peptide; 7 AA.
 XX XX
 AC XX ABP48802;
 XX XX
 XX XX 28-AUG-2002 (first entry)
 DT XX
 XX XX

DE Zinc finger protein related peptide motif SEQ ID NO:944.
 KW Zinc finger protein; ZFP; DNA binding protein; zinc finger.
 XX

OS Homo sapiens.
 OS Synthetic.

XX WO200242459-A2.
 XX

XX 30-MAY-2002.
 XX

XX 30-MAY-2002.
 XX

XX 20-NOV-2001; 2001WO-US43438.
 XX

XX 20-NOV-2000; 2000US-0716637.
 XX

XX (SANG-) SANGAMO BIOSCIENCES INC.
 XX

XX Liu Q;
 XX

XX WPI; 2002-500284/53.
 XX

XX New zinc finger protein that binds to target site, useful in studying
 XX gene function and for human therapeutics and plant engineering.
 XX comprises first, second and third zinc fingers, ordered from N- to
 XX C-terminus

XX Example 1; Page 42; 81pp; English.
 XX

XX The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target subsite. Also described are: (1) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC it binds to the S1 target subsite, selecting the F1 zinc finger such that
 CC that it binds to the S2 target subsite, and selecting the F2 zinc finger such
 CC that it binds to the S3 target subsite, thus designing (I)
 CC finger such that it binds to the S3 target subsite, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target subsites having the nucleotide G in the 5'-most position of the
 CC subsite. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determine the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences. (I) has improved affinity
 CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.
 XX

SQ Sequence 7 AA;

Query Match 91.7%; Score 33; DB 23; Length 7;
 Best Local Similarity 85.7%; Pred. No. 7.8e+05;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLR 7
 |||||:|
 Db 1 TSGHLR 7

RESULT 9
 ABP49078

ID ABP49078 standard; Peptide; 7 AA.

XX AC ABP49078;

XX 28-AUG-2002 (first entry)

XX Zinc finger protein related peptide motif SEQ ID NO:1036.

XX Zinc finger protein; ZFP; DNA binding protein; zinc finger.

XX Homo sapiens.
 OS Synthetic.

XX WO200242459-A2.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US43438.

XX 20-NOV-2000; 2000US-0716637.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Liu Q;

XX WPI; 2002-500284/53.

XX New zinc finger protein that binds to target site, useful in studying
 XX gene function and for human therapeutics and plant engineering.
 XX comprises first, second and third zinc fingers, ordered from N- to
 XX C-terminus

XX Example 1; Page 45; 81pp; English.

XX The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target subsite. Also described are: (1) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC it binds to the S1 target subsite, selecting the F1 zinc finger such that
 CC that it binds to the S2 target subsite, and selecting the F2 zinc finger such
 CC that it binds to the S3 target subsite, thus designing (I)
 CC finger such that it binds to the S3 target subsite, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target subsites having the nucleotide G in the 5'-most position of the
 CC subsite. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determine the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences. (I) has improved affinity
 CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.
 XX

SQ Sequence 7 AA;

Query Match 91.7%; Score 33; DB 23; Length 7;
 Best Local Similarity 85.7%; Pred. No. 7.8e+05;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLR 7
 |||||:|
 Db 1 TSGHLR 7

RESULT 10
 ABP49085

ID ABP49085 standard; Peptide; 7 AA.

XX AC ABP49085;

XX 28-AUG-2002 (first entry)

XX Zinc finger protein related peptide motif SEQ ID NO:1243.

XX Zinc finger protein; ZFP; DNA binding protein; zinc finger.

XX Homo sapiens.

OS Synthetic.

XX WO200242459-A2.
 XX 30-MAY-2002.
 PD 20-NOV-2001; 2001WO-US43438.
 XX 20-NOV-2000; 2000US-0716637.
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 PA Liu Q;
 PI WPI; 2002-500284/53.
 XX New zinc finger protein that binds to target site, useful in studying
 XX gene function and for human therapeutics and plant engineering,
 PT comprises first, second and third zinc fingers, ordered from N- to
 PT C-terminus
 PS Example 1; Page 45; 81pp; English.
 XX The present invention describes a zinc finger protein (I) that binds to
 XX a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target subsite. Also described are: (1) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target subsite, selecting the F2 zinc finger such
 CC that it binds to the S2 target subsite, and selecting the F3 zinc
 CC finger such that it binds to the S3 target subsite, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target subsites having the nucleotide G in the 5'-most position of the
 CC subsite. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determine the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. AB071213 to AB072214 and ABP48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.
 XX Sequence 7 AA;
 SQ Query Match 91.7%; Score 33; DB 23; Length 7;
 Best Local Similarity 85.7%; Pred. No. 7.8e+05;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TSGHLISR 7
 Db 1 TSGHLTR 7
 RESULT 11
 ABP49088
 ID ABP49088 standard; Peptide; 7 AA.
 XX AC ABP49088;
 XX 28-AUG-2002 (first entry)
 DT Zinc finger protein related peptide motif SEQ ID NO:1244.
 DE Zinc finger protein; ZFP; DNA binding protein; zinc finger.
 KW Homo sapiens.
 XX OS Synthetic.
 XX WO200242459-A2.
 XX 30-MAY-2002.
 XX 20-NOV-2001; 2001WO-US43438.

PD 30-MAY-2002.
 XX 20-NOV-2001; 2001WO-US43438.
 XX 20-NOV-2000; 2000US-0716637.
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 PA Liu Q;
 PI WPI; 2002-500284/53.
 XX New zinc finger protein that binds to target site, useful in studying
 XX gene function and for human therapeutics and plant engineering,
 PT comprises first, second and third zinc fingers, ordered from N- to
 PT C-terminus
 PS Example 1; Page 45; 81pp; English.
 XX The present invention describes a zinc finger protein (I) that binds to
 XX a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target subsite. Also described are: (1) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target subsite, selecting the F2 zinc finger such
 CC that it binds to the S2 target subsite, and selecting the F3 zinc
 CC finger such that it binds to the S3 target subsite, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target subsites having the nucleotide G in the 5'-most position of the
 CC subsite. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determine the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. AB071213 to AB072214 and ABP48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.
 XX Sequence 7 AA;
 SQ Query Match 91.7%; Score 33; DB 23; Length 7;
 Best Local Similarity 85.7%; Pred. No. 7.8e+05;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TSGHLISR 7
 Db 1 TSGHLTR 7
 RESULT 12
 ABP49102
 ID ABP49102 standard; Peptide; 7 AA.
 XX AC ABP49102;
 XX 28-AUG-2002 (first entry)
 DT Zinc finger protein related peptide motif SEQ ID NO:1044.
 DE Zinc finger protein; ZFP; DNA binding protein; zinc finger.
 KW Homo sapiens.
 XX OS Synthetic.
 XX WO200242459-A2.
 XX 30-MAY-2002.
 XX 20-NOV-2001; 2001WO-US43438.

XX 20-NOV-2000; 2000US-0716637.
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX Liu Q;
XX WPI; 2002-500284/53.
XX New zinc finger protein that binds to target site, useful in studying
PT gene function and for human therapeutics and plant engineering,
PT comprises first, second and third zinc fingers, ordered from N- to
PT C-terminus
XX
XX Example 1; Page 45; 81pp; English.
XX The present invention describes a zinc finger protein (I) that binds to
CC a target site, comprising a first (F1), a second (F2), and a third (F3)
CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
CC and a third (S3) target subsite. Also described are: (1) a polypeptide
CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
CC it binds to the S1 target subsite, selecting the F2 zinc finger such that
CC it binds to the S2 target subsite, selecting the F3 zinc finger such that
CC that it binds to the S3 target subsite, and selecting the F1 zinc finger such
CC that it binds to a target site. (I) is useful for recognition of triplet
CC target subsites having the nucleotide G in the 5'-most position of the
CC subsite. (I) is useful in studying gene function, and for human
CC therapeutics and plant engineering. (I), (II) or (III) is useful in
CC therapeutic methods to modulate the expression of a target region within
CC a subject, in diagnostic methods for sequence specific detection of
CC target nucleic acid in a sample, and in assays to determine the
CC phenotype and function of gene expression. (I) has improved affinity
CC and specificity for their target sequences, as well as enhanced
CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.
XX
XX Sequence 7 AA;
SQ
Query Match 91.7%; Score 33; DB 23; Length 7;
Best Local Similarity 85.7%; Pred. NO. 7.8e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TSGHLR 7
DB 1 TSGHLR 7
RESULT 13
ABP48958
ID ABP48958 standard; Peptide; 7 AA.
XX ABP48958;
XX
XX 28-AUG-2002 (first entry)
DT
DE Zinc finger protein related peptide motif SEQ ID NO:996.
XX
XX Zinc finger protein; ZFP; DNA binding protein; zinc finger.
XX Homo sapiens.
OS Synthetic.
XX
XX WO200242459-A2.
PN
XX
XX 30-MAY-2002.
PD
XX
XX 20-NOV-2001; 2001WO-US43438.
PF
XX
XX 20-NOV-2000; 2000US-0716637.
PR
XX

(SANG-) SANGAMO BIOSCIENCES INC.
Liu Q;
WPI; 2002-500284/53.
New zinc finger protein that binds to target site, useful in studying
gene function and for human therapeutics and plant engineering,
comprises first, second and third zinc fingers, ordered from N- to
C-terminus
Example 1; Page 44; 81pp; English.
The present invention describes a zinc finger protein (I) that binds to
a target site, comprising a first (F1), a second (F2), and a third (F3)
zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
target site comprises, in 3'-5' direction, a first (S1), a second (S2),
and a third (S3) target subsite. Also described are: (1) a polypeptide
(II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
(3) designing (M) (I) involves selecting the F1 zinc finger such that
it binds to the S1 target subsite, selecting the F2 zinc finger such that
it binds to the S2 target subsite, and selecting the F3 zinc finger such
that it binds to the S3 target subsite, thus designing (I)
that binds to a target site. (I) is useful for recognition of triplet
target subsites having the nucleotide G in the 5'-most position of the
subsite. (I) is useful in studying gene function, and for human
therapeutics and plant engineering. (I), (II) or (III) is useful in
therapeutic methods to modulate the expression of a target region within
a subject, in diagnostic methods for sequence specific detection of
target nucleic acid in a sample, and in assays to determine the
phenotype and function of gene expression. (I) has improved affinity
and specificity for their target sequences, as well as enhanced
biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
represent DNA target sequences and zinc finger peptides which are given
in the exemplification of the present invention.
XX
XX Sequence 7 AA;
SQ
Query Match 88.9%; Score 32; DB 23; Length 7;
Best Local Similarity 85.7%; Pred. NO. 7.8e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TSGHLR 7
DB 1 TSGHLR 7
RESULT 14
AAG14909
ID AAG14909 standard; Protein; 284 AA.
XX
XX AAG14909;
XX
XX 17-OCT-2000 (first entry)
DT
DE Arabidopsis thaliana protein fragment SEQ ID NO: 14948.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR
XX 05-MAR-1999; 99US-0123180.
PR
XX 09-MAR-1999; 99US-0123548.
PR
XX 23-MAR-1999; 99US-0125788.
PR

PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142920.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147036.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.

PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.9%; Score 32; DB 21; Length 284;

Best Local Similarity 85.7%; Pred. No. 1e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TSGHLSR 7

Db 234 TSGHLQR 240

RESULT 15

AAG45427 AAG45427 standard; Protein; 285 AA.

ID XX

AC AAG45427;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 57031.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 29-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 18-APR-1999; 99US-0128714.

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0158458.
PR 29-SEP-1999; 99US-0158596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.9%; Score 32; DB 21; Length 285;
Best Local Similarity 85.7%; Pred. No. 1a+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
| | | | |
DB 235 TSGHLQR 241

Search completed: July 15, 2003, 12:53:01
Job time : 33.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:50:16 ; Search time 23.6667 Seconds
(without alignments)
60.944 Million cell updates/sec

Title: US-10-006-069a-247

Perfect score: 36

Sequence: 1 TSGHLSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	33	91.7	86	2	O87366
2	33	91.7	310	3	O9C410
3	32	88.9	155	11	O8VIB4
4	32	88.9	285	10	O9MI42
5	32	88.9	305	2	O8VPM4
6	32	88.9	368	10	O9LWY8
7	32	88.9	542	4	O9NWV6
8	32	88.9	580	4	O96C45
9	32	88.9	910	11	O9D4H6
10	31	86.1	116	5	O46363
11	31	86.1	131	11	O9DIL2
12	31	86.1	139	11	O9JW97
13	31	86.1	153	10	O9ASA1
14	31	86.1	265	10	O9LQ29
15	31	86.1	286	5	O17595
16	31	86.1	378	5	O45110

17	31	86.1	379	13	O9DG05
18	31	86.1	423	12	O57192
19	31	86.1	423	12	O9JFC4
20	31	86.1	423	12	O8V512
21	31	86.1	423	12	O8V2U4
22	31	86.1	423	12	O8QWY4
23	31	86.1	424	13	O9DG06
24	31	86.1	457	10	O94LG3
25	31	86.1	496	10	O9SJK1
26	31	86.1	503	11	O91286
27	31	86.1	525	6	O9BGR3
28	31	86.1	531	16	O32138
29	31	86.1	639	16	O9ZC51
30	31	86.1	752	4	O13597
31	31	86.1	762	10	O22168
32	31	86.1	765	10	O8RY37
33	31	86.1	765	11	O9D2N7
34	31	86.1	1124	10	O8RXC2
35	31	86.1	2658	5	O9GRL9
36	30	83.3	189	5	O9N8S7
37	30	83.3	361	2	O60100
38	30	83.3	400	5	O9XTB4
39	30	83.3	615	3	O9HEX0
40	30	83.3	680	16	O9C1H7
41	30	83.3	783	4	O96CC8
42	30	83.3	1644	16	O8XC72
43	30	83.3	1752	10	O9LX60
44	30	83.3	2301	5	O9N601
45	30	83.3	3589	2	O69246

ALIGNMENTS

RESULT 1

O87366 PRELIMINARY; PRT; 86 AA.

AC O87366;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 10.2 kDa protein.
OS Staphylococcus aureus.
OG Plasmid pSV41.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
[1]

RP SEQUENCE FROM N.A.
RX MEDLINE=98389645; PubMed=9721269;
RA Berg T., Firth N., Apisiridej S., Hettiaratchi A., Leelaporn A.,
RA Skurray R.A.;
RT "Complete nucleotide sequence of pSK41: evolution of staphylococcal
RT conjugative multiresistance plasmids.";
RL J. Bacteriol. 180:4350-4359(1998).
DR EMBL; AF051917; AAC61943.1;
DR InterPro; IPR001387; HTH_3.
DR Pfam; PF01381; HTH_3; 1.
DR SMART; SM00530; HTH_XRE; 1.
KW Hypothetical protein; Plasmid.

SQ SEQUENCE 86 AA; 10208 MW; F38D9F7AD2DD4450 CRC64;
Query Match 91.7%; Score 33; DB 2; Length 86;
Best Local Similarity 85.7%; Pred. No. 7.8; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 0;

Qy 1 TSGHLSR 7

Db 38 TAGHLSR 44

RESULT 2

O9C410

```

ID Q9C410 PRELIMINARY; PRT; 310 AA.
AC Q9C410;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Nrgl.
GN NRG1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10261;
RX MEDLINE=21423454; PubMed=11532938;
RA Murad A.M.A., Leng P., Straffon M., Wishart J., Macaskill S.,
RA MacCallum D., Schnell N., Talibi D., Marschal D., Tekala F.,
RA d'Enfert C., Gaillardin C., Odds F.C., Brown A.J.P.;
RT 'NRG1 represses yeast-hypha morphogenesis and hypha-specific gene
RT expression in Candida albicans.';
RL EMBO J. 20:4742-4752(2001).
DR EMBL; AF321521; AAK09366.1; -.
DR HSSP; P07248; IARD.
DR InterPro; IPR000822; znf_C2H2.
DR Pfam; PF00096; zf_C2H2; 2.
DR ProDom; PD000003; znf_C2H2; 1.
DR SMART; SM00355; znf_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 310 AA; 34301 MW; 1E108C0998A51284 CRC64;

Query Match 91.7%; Score 33; DB 3; Length 310;
Best Local Similarity 85.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
DB 239 TSGHLAR 245

RESULT 3
Q8VB4 PRELIMINARY; PRT; 155 AA.
ID Q8VB4;
AC Q8VB4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Dystrophin-like protein (Fragment).
GN DYSLIKE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RA Takeuchi T.;
RT "Mouse dystrophin-like gene.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RA Takeuchi T.;
RT "New dystrophin-like protein in heart.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB063290; BAB79447.1; -.
DR InterPro; IPR000433; znf_22.
DR Pfam; PF00569; Z2; 1.
DR SMART; SM00291; znf_22; 1.
DR PROSITE; PS01357; ZF_22_1; UNKNOWN_1.
DR PROSITE; PS0135; ZF_22_2; 1.
FT NON_TER 1

```

```

SQ SEQUENCE 155 AA; 17998 MW; 30AFE0E1BE045D42 CRC64;

Query Match 88.9%; Score 32; DB 11; Length 155;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
DB 108 TEGHLSR 114

RESULT 4
Q9M142 PRELIMINARY; PRT; 285 AA.
ID Q9M142;
AC Q9M142;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative myb-related DNA-binding protein.
GN AT4G01280.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
DR EMBL; AL161491; CAB80937.1; -.
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF00249; myb_DNA-binding; 1.
DR SMART; SM00395; SANT; 1.
DR PROSITE; PS50090; MYB_3; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 285 AA; 31930 MW; 0666490377D0D700 CRC64;

Query Match 88.9%; Score 32; DB 10; Length 285;
Best Local Similarity 85.7%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
DB 235 TSGHLQR 241

RESULT 5
Q8VPM4 PRELIMINARY; PRT; 305 AA.
ID Q8VPM4;
AC Q8VPM4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative glycine-rich cell wall protein.
OS Micrococcus sp. 28.
OG Plasmid pSD10.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcales; Micrococcaceae; Micrococcus.
OX NCBI_TaxID=161213;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=28;
RA Zhong Z., Caspi R., Mincer T., Helinski D., Knauf V., Boardman K.,
RA Wilkinson J.E., Shea T., DeLoughery C., Toukdarian A.;
RT "A 50 kb plasmid rich in mobile gene sequences isolated from a marine
RT Micrococcus.";

```

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY034092; AA62523.1; -
 KW Plasmid.

SQ SEQUENCE 305 AA; 3008 MW; 8AAB816F9FF392D5 CRC64;

Query Match 88.9%; Score 32; DB 2; Length 305;
 Best Local Similarity 85.7%; Pred. No. 47;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
 | | | | |
 Db 184 TGGHLSR 190

RESULT 6

ID Q9LWY8 PRELIMINARY; PRT; 368 AA.
 AC Q9LWY8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Similar to Arabidopsis thaliana DNA chromosome 4.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartidae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
 clone: P0644B06";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP001129; BAA90626.1; -
 DR InterPro; IPR000620; DUF6.
 DR Pfam; PF00892; DUF6; 1.
 SQ SEQUENCE 368 AA; 40102 MW; 056EF1F9059C0470 CRC64;

Query Match 88.9%; Score 32; DB 10; Length 368;
 Best Local Similarity 85.7%; Pred. No. 58;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
 | | | | |
 Db 19 TSGHLR 25

RESULT 7

ID Q9NWV6 PRELIMINARY; PRT; 542 AA.
 AC Q9NWV6;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE CDNA FLJ20574 f1s, clone REC01035.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K.,
 RA Hirao M., Ohmori Y., Ota T., Suzuki Y., Ohayashi M., Nishi T.,
 RA Shibahara T., Tanaka T., Nakamura Y., Isegai T., Sugano S.;
 RT "NEDO human cDNA sequencing project";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK00581; BAA91270.1; -
 DR InterPro; IPR000719; Euk_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 2.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM;
 KW ATP-binding; Transferase.

SQ SEQUENCE 542 AA; 60457 MW; 1988711A61EA1AAF CRC64;

Query Match 88.9%; Score 32; DB 4; Length 542;
 Best Local Similarity 85.7%; Pred. No. 85;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
 | | | | |
 Db 396 TSGHLSQ 402

RESULT 8

ID Q96C45 PRELIMINARY; PRT; 580 AA.
 AC Q96C45;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical 65.0 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LYMPH;
 RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC014794; AAH14794.1; -
 DR InterPro; IPR00719; Euk_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 2.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM;
 KW ATP-binding; Hypothetical protein; Transferase.
 SQ SEQUENCE 580 AA; 64973 MW; 61426673F6B2CEAD CRC64;

Query Match 88.9%; Score 32; DB 4; Length 580;
 Best Local Similarity 85.7%; Pred. No. 92;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
 | | | | |
 Db 396 TSGHLSQ 402

RESULT 9

ID Q9D4H6 PRELIMINARY; PRT; 910 AA.
 AC Q9D4H6;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE 4932415A06Rik protein.
 GN 4932415A06Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

```

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK016524; BAB30285.1; -.
DR HSSP; P24941; 1HCL.
DR MGD; MGI:1921622; 4932415A06Rik.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.
SQ SEQUENCE 910 AA; 101937 MW; A602520F381A0ECD CRC64;

Query Match 88.9%; Score 32; DB 11; Length 910;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
| | | | |
DB 396 TSGHLSQ 402

RESULT 10
O46363
ID O46363 PRELIMINARY; PRT; 116 AA.
AC O46363
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Universal minicircle sequence binding protein.
GN UMSBP2 OR UMSBP1.
OS Crithidia fasciculata.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Crithidia.
OX NCBI_TaxID=5636;
RN [1]
RP SEQUENCE FROM N.A.
RA Tzfat Y., Shlomai J.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF045246; AAC32814.1; -.
DR EMBL; AF045246; AAC32813.1; -.
DR InterPro; IPR001878; znf_CCHC.
DR Pfam; PF00098; zf-CCHC; 5.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; znf_C2HC; 5.
SQ SEQUENCE 116 AA; 12684 MW; F4BD053987585B1A CRC64;

Query Match 86.1%; Score 31; DB 5; Length 116;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGHLSR 7
| | | | |
DB 83 SGHLSR 88

RESULT 11
Q9D1L2
ID Q9D1L2 PRELIMINARY; PRT; 131 AA.
AC Q9D1L2
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 1110003008Rik protein.

```

```

GN 1110003008Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK003388; BAB22757.1; -.
DR MGD; MGI:1923541; 1110003008Rik.
SQ SEQUENCE 131 AA; 14740 MW; F5B859EEDC98D4B8 CRC64;

Query Match 86.1%; Score 31; DB 11; Length 131;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
| | | | |
DB 18 TSGHLSR 24

RESULT 12
Q9JM97
ID Q9JM97 PRELIMINARY; PRT; 139 AA.
AC Q9JM97
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Estrogen receptor beta (Fragment).
GN ESR2 OR ERB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B/C;
RX MEDLINE=21311378; PubMed=11418189;
RA Ishibashi O., Kawashima H.;
RT "Cloning and characterization of the functional promoter of mouse
RT estrogen receptor beta gene.";
RL Biochim. Biophys. Acta 1519:223-229(2001).
DR EMBL; AB034957; BAA92313.1; -.
DR MGD; MGI:109392; Esr2.
KW Receptor.
FT NON_TER 139 139
SQ SEQUENCE 139 AA; 15215 MW; 4703615AE8593DAB CRC64;

Query Match 86.1%; Score 31; DB 11; Length 139;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

QY 1 TSGHLS 6
    |||||
Db 101 TSGHLS 106

RESULT 13
Q9ASAI PRELIMINARY; PRT; 153 AA.
AC Q9ASAI:
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE p0707D10.26 protein.
GN p0707D10.26.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:p0707D10."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002910; BAB40168.1;
SQ SEQUENCE 153 AA; 17321 MW; 04DADB774FAB3B9D CRC64;

Query Match 86.1%; Score 31; DB 10; Length 153;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLS 7
    |||||
Db 115 TSGHLR 121

RESULT 14
Q9LQ29 PRELIMINARY; PRT; 265 AA.
AC Q9LQ29:
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE F10A5.22
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC006434; AAF87117.1;
DR InterPro: IPR001878; Znf_CCHC.

DR Pfam: PF00098; zf-CCHC; 8.
DR SMART: SM00343; Znf_C2HC; 8.
SQ SEQUENCE 265 AA; 29334 MW; BCD90D691F82762 CRC64;

Query Match 86.1%; Score 31; DB 10; Length 265;
Best Local Similarity 71.4%; Pred. No. 69;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 7
    |||||
Db 172 TSGHIAR 178

RESULT 15
O17595 PRELIMINARY; PRT; 286 AA.
AC O17595:
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE C25D7.2 protein.
GN C25D7.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Alnscough R.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z81039; CAB02773.1;
DR InterPro: IPR000535; MSP_domain.
DR Pfam: PF00635; MSP_domain; 1.
DR PRINTS: PR01608; BACINVASINC.
SQ SEQUENCE 286 AA; 30470 MW; 0A28C9170478D85F CRC64;

Query Match 86.1%; Score 31; DB 5; Length 286;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGHLSR 7
    |||||
Db 55 SGHLSR 60

Search completed: July 15, 2003, 12:54:51
Job time : 25.6667 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:51:16 ; Search time 10.6667 Seconds
(without alignments)
19.309 Million cell updates/sec

Title: US-10-006-069A-247
Perfect score: 36
Sequence: 1 TSGHLSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCRU5.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	86.1	173	4	US-09-228-986-104
2	31	86.1	484	2	US-08-836-620A-13
3	31	86.1	484	2	US-08-836-620A-14
4	31	86.1	485	2	US-08-836-620A-2
5	31	86.1	485	2	US-08-836-620A-5
6	30	83.3	7	4	US-09-173-941-12
7	28	77.8	40	1	US-07-743-518-18
8	28	77.8	228	1	US-08-447-591-4
9	28	77.8	228	1	US-08-450-943-4
10	28	77.8	228	1	US-08-059-031-4
11	28	77.8	228	2	US-08-450-942-4
12	28	77.8	228	5	PCT-US94-05090-4
13	28	77.8	272	4	US-09-069-023-24
14	28	77.8	296	4	US-09-134-001C-3190
15	28	77.8	391	3	US-08-968-563-11
16	28	77.8	391	4	US-08-969-683A-11
17	28	77.8	391	4	US-09-297-928-7
18	28	77.8	460	3	US-08-935-263-4
19	28	77.8	460	3	US-09-594-185-4
20	28	77.8	466	3	US-08-526-136-13
21	28	77.8	486	2	US-08-911-445-18
22	28	77.8	486	4	US-09-182-983-18
23	28	77.8	503	3	US-08-526-136-2
24	28	77.8	505	3	US-08-526-136-4
25	28	77.8	631	4	US-08-971-188-8
26	28	77.8	1832	3	US-09-335-409-4
27	28	77.8	1832	4	US-09-568-102-4

28	77.8	1832	4	US-09-567-969-4	Sequence 4, Appli
29	77.8	1832	4	US-09-568-480-4	Sequence 4, Appli
30	77.8	1832	4	US-09-568-486-4	Sequence 4, Appli
31	77.8	1832	4	US-09-568-472-4	Sequence 4, Appli
32	77.8	1832	4	US-09-567-899-4	Sequence 24, Appli
33	27	75.0	53	2	US-08-469-537A-24
34	27	75.0	54	2	US-08-456-647B-50
35	27	75.0	54	2	US-08-237-401A-50
36	27	75.0	76	2	US-08-469-537A-45
37	27	75.0	171	2	US-08-609-049A-19
38	27	75.0	171	4	US-09-170-996-19
39	27	75.0	175	4	US-08-887-534A-14
40	27	75.0	244	1	US-08-361-395-1
41	27	75.0	269	4	US-09-134-001C-5197
42	27	75.0	342	3	US-09-136-073-2
43	27	75.0	342	4	US-09-457-024A-2
44	27	75.0	358	4	US-09-071-035-492
45	27	75.0	388	4	US-09-071-035-490

ALIGNMENTS

RESULT 1
US-09-228-986-104
; Sequence 104, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-104

Query Match 86.1%; Score 31; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TSGHLS 6
Db 44 TSGHLS 49

RESULT 2
US-08-836-620A-13
; Sequence 13, Application US/08836620A
; Patent No. 5958710
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Orphan receptor
; NUMBER OF SEQUENCES: 19
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,620A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/03933
; FILING DATE:
; APPLICATION NUMBER: GB 9518272.1
; FILING DATE: 08-SEP-1995

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9605550.4
;; FILING DATE: 15-MAR-1996
;; PRIOR APPLICATION NUMBER: GB 9607532.0
;; FILING DATE: 11-APR-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9609576.5
;; FILING DATE: 08-MAY-1996
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 484 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; ORGANISM: Rattus rattus
US-08-836-620A-13

Query Match 86.1%; Score 31; DB 2; Length 484;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
Db 37 TSGHLS 42

RESULT 3
US-08-836-620A-14
; Sequence 14, Application US/08836620A
; Patent No. 5958710
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Orphan receptor
; NUMBER OF SEQUENCES: 19
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,620A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/03933
; FILING DATE:
; APPLICATION NUMBER: GB 9518272.1
; FILING DATE: 08-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9605550.4
; FILING DATE: 15-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9607532.0
; FILING DATE: 11-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9609576.5
; FILING DATE: 08-MAY-1996
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 484 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; ORGANISM: Mus musculus
US-08-836-620A-14

Query Match 86.1%; Score 31; DB 2; Length 484;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
Db 37 TSGHLS 42

RESULT 4
US-08-836-620A-2
; Sequence 2, Application US/08836620A
; Patent No. 5958710
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Orphan receptor
; NUMBER OF SEQUENCES: 19
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,620A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/03933
; FILING DATE:
; APPLICATION NUMBER: GB 9518272.1
; FILING DATE: 08-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9605550.4
; FILING DATE: 15-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9607532.0
; FILING DATE: 11-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9609576.5
; FILING DATE: 08-MAY-1996
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 485 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; ORGANISM: Rattus rattus
US-08-836-620A-2

Query Match 86.1%; Score 31; DB 2; Length 485;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
Db 37 TSGHLS 42

RESULT 5
US-08-836-620A-5
; Sequence 5, Application US/08836620A
; Patent No. 5958710
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Orphan receptor
; NUMBER OF SEQUENCES: 19
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,620A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/03933
; FILING DATE:
; APPLICATION NUMBER: GB 9518272.1
; FILING DATE: 08-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9605550.4

Query Match 86.1%; Score 31; DB 2; Length 484;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
Db 37 TSGHLS 42

; FILING DATE: 15-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9607532.0
; FILING DATE: 11-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9609576.5
; FILING DATE: 08-MAY-1996
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 485 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; ORGANISM: Mus musculus
; US-08-836-620A-5

Query Match 86.1%; Score 31; DB 2; Length 485;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TSGHLS 6
Db 37 TSGHLS 42

RESULT 6

US-09-173-941-12
; Sequence 12, Application US/09173941
; Patent No. 6140081

; GENERAL INFORMATION:
; APPLICANT: BARBAS, Carlos F.
; TITLE OF INVENTION: ZINC FINGER BINDING DOMAINS FOR GNN
; FILE REFERENCE: NOV0081S
; CURRENT APPLICATION NUMBER: US/09/173,941
; CURRENT FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: nucleotide
; OTHER INFORMATION: codon binding sequence

US-09-173-941-12

Query Match 83.3%; Score 30; DB 4; Length 7;
Best Local Similarity 85.7%; Pred. No. 1.9e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TSGHLS 7
Db 1 TSGHLV 7

RESULT 7

US-07-743-518-18
; Sequence 18, Application US/07743518
; Patent No. 5397696

; GENERAL INFORMATION:
; APPLICANT: YANAGIHARA, RICHARD
; APPLICANT: NERURKAR, VIVEK R.
; APPLICANT: JENKINS, CAROL
; APPLICANT: MILLER, MARK
; APPLICANT: GARUTO, RALPH M.
; TITLE OF INVENTION: PAPOA NEW GUINEA HUMAN T-LYMPHOTROPIC
; TITLE OF INVENTION: VIRUS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.

; COUNTRY: U.S.A.
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/743,518
; FILING DATE: 19910812
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: KTS/5683/84699/SAP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-743-518-18

Query Match 77.8%; Score 28; DB 1; Length 40;
Best Local Similarity 83.3%; Pred. No. 21;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SGHLSR 7
Db 27 TGHLSR 32

RESULT 8

US-08-447-591-4
; Sequence 4, Application US/08447591
; Patent No. 5591440

; GENERAL INFORMATION:
; APPLICANT: CARMAN, WILLIAM
; APPLICANT: DECKER, RICHARD H
; APPLICANT: WALLACE, LESLEY
; APPLICANT: MIMMS, LARRY T
; APPLICANT: SOLOMON, LARRY R
; TITLE OF INVENTION: HEPATITIS B VIRUS MUTANTS, REAGENTS AND METHODS FOR DETECTI
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: ONE ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,591
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/059,031
; FILING DATE: 07-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: POREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5347.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365

TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-447-591-4

Query Match 77.8%; Score 28; DB 1; Length 228;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
|||||
DB 5 TSGHLT 10

RESULT 9

US-08-450-943-4
; Sequence 4, Application US/08450943
; Patent No. 593825
; GENERAL INFORMATION:
; APPLICANT: CARMAN, WILLIAM
; APPLICANT: DECKER, RICHARD H
; APPLICANT: WALLIS, LESLEY
; APPLICANT: MIMMS, LARRY T
; APPLICANT: SOLOMON, LARRY R
; TITLE OF INVENTION: HEPATITIS B VIRUS MUTANTS, REAGENTS AND METHODS FOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: ONE ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,943
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5347.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-450-943-4

Query Match 77.8%; Score 28; DB 1; Length 228;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
|||||
DB 5 TSGHLT 10

RESULT 10
US-08-059-031-4
; Sequence 4, Application US/08059031
; Patent No. 595739
; GENERAL INFORMATION:
; APPLICANT: CARMAN, WILLIAM
; APPLICANT: DECKER, RICHARD H
; APPLICANT: WALLACE, LESLEY
; APPLICANT: MIMMS, LARRY T
; APPLICANT: SOLOMON, LARRY R
; TITLE OF INVENTION: HEPATITIS B VIRUS MUTANTS, REAGENTS AND METHODS FOR DETECTI
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: ONE ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/059,031
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5347.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-059-031-4

Query Match 77.8%; Score 28; DB 1; Length 228;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
|||||
DB 5 TSGHLT 10

RESULT 11

US-08-450-942-4
; Sequence 4, Application US/08450942
; Patent No. 592512
; GENERAL INFORMATION:
; APPLICANT: CARMAN, WILLIAM
; APPLICANT: DECKER, RICHARD H
; APPLICANT: WALLACE, LESLEY
; APPLICANT: MIMMS, LARRY T
; APPLICANT: SOLOMON, LARRY R
; TITLE OF INVENTION: HEPATITIS B VIRUS MUTANTS, REAGENTS AND METHODS FOR DETECTI
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: ONE ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,942
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5347.US.01
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-450-942-4

Query Match 77.8%; Score 28; DB 2; Length 228;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
| | | | |
Db 5 TSGHLT 10

RESULT 12
PCT-US94-05090-4
; Sequence 4, Application PC/TUS9405090
; GENERAL INFORMATION:
; APPLICANT: CARMAN, WILLIAM
; APPLICANT: DECKER, RICHARD H
; APPLICANT: WALLACE, LESLEY
; APPLICANT: MINNS, LARRY T
; APPLICANT: SOLOMON, LARRY R
; TITLE OF INVENTION: HEPATITIS B VIRUS MUTANTS, REAGENTS AND METHODS FOR DETECTION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: ONE ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05090
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5347.US.01
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; PCT-US94-05090-4

Query Match 77.8%; Score 28; DB 5; Length 228;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
| | | | |
Db 5 TSGHLT 10

RESULT 13
US-09-069-023-24
; Sequence 24, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-069-023-24

Query Match 77.8%; Score 28; DB 4; Length 272;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
| | | | |
Db 58 TSGHVS 63

RESULT 14
US-09-134-001C-3190
; Sequence 3190, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3190
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-3190

Query Match 77.8%; Score 28; DB 4; Length 296;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
| | | | |
Db 93 TEGHLKR 99

Job time : 11.6667 secs

RESULT 15
US-08-968-563-11
; Sequence 11, Application US/08968563
; Patent No. 6013494
; GENERAL INFORMATION:
; APPLICANT: CHARLES E. NAKAMURA
; APPLICANT: ANTHONY A. GATENBY
; APPLICANT: AMY (KUANG-HUA) HSU
; APPLICANT: RICHARD D. LA REAU
; APPLICANT: SHARON L. HAYNIE
; APPLICANT: MARIA DIAZ-TORRES
; APPLICANT: DONALD E. TRIMBUR
; APPLICANT: GREGORY M. WHITED
; APPLICANT: VASANTHA NAGARAJAN
; APPLICANT: MARK S. PAYNE
; APPLICANT: STEPHEN K. PICATAGGIO
; APPLICANT: RAMESCH V. NAIR
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; TITLE OF INVENTION: PRODUCTION OF 1,3-PROPANEDIOL
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; ADDRESSEE: GENENCOR INTERNATIONAL, INC.
; STREET: 4 CAMBRIDGE PLACE
; STREET: 1870 SOUTH WINTON ROAD
; CITY: ROCHESTER
; STATE: NEW YORK
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 INCH DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,563
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,601
; FILING DATE: NOVEMBER 13, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9982
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: GFDI
US-08-968-563-11

Query Match 77.8%; Score 28; DB 3; Length 391;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLS 6
DB 11 TSGHLN 16

Search completed: July 15, 2003, 12:56:10

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 12:53:06 ; Search time 16.3333 Seconds
(without alignments)
49.894 Million cell updates/sec

Title: US-10-006-069A-247
Perfect score: 36
Sequence: 1 TSGHLSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	7	9	US-10-006-069A-247
2	36	100.0	7	9	Sequence 247, App
3	36	100.0	7	9	Sequence 247, App
4	36	100.0	7	9	Sequence 1201, Ap
5	36	100.0	7	10	Sequence 1201, Ap
6	33	91.7	7	9	Sequence 245, App
7	33	91.7	7	9	Sequence 248, App
8	33	91.7	7	9	Sequence 245, App
9	33	91.7	7	9	Sequence 248, App
10	33	91.7	7	9	Sequence 393, App
11	33	91.7	7	9	Sequence 738, App
12	33	91.7	7	9	Sequence 944, App
13	33	91.7	7	9	Sequence 1036, Ap
14	33	91.7	7	9	Sequence 1044, Ap
15	33	91.7	7	9	Sequence 1147, Ap
16	33	91.7	7	9	Sequence 1243, Ap
17	33	91.7	7	9	Sequence 1244, Ap
18	33	91.7	7	9	Sequence 393, App
19	33	91.7	7	9	Sequence 738, App

20	33	91.7	7	9	US-09-989-994-944	Sequence 944, App
21	33	91.7	7	9	US-09-989-994-1036	Sequence 1036, Ap
22	33	91.7	7	9	US-09-989-994-1044	Sequence 1044, Ap
23	33	91.7	7	9	US-09-989-994-1147	Sequence 1147, Ap
24	33	91.7	7	9	US-09-989-994-1243	Sequence 1243, Ap
25	33	91.7	7	9	US-09-989-994-1244	Sequence 1244, Ap
26	33	91.7	7	10	US-09-989-789-393	Sequence 393, App
27	33	91.7	7	10	US-09-989-789-738	Sequence 738, App
28	33	91.7	7	10	US-09-989-789-944	Sequence 944, App
29	33	91.7	7	10	US-09-989-789-1036	Sequence 1036, Ap
30	33	91.7	7	10	US-09-989-789-1044	Sequence 1044, Ap
31	33	91.7	7	10	US-09-989-789-1147	Sequence 1147, Ap
32	33	91.7	7	10	US-09-989-789-1243	Sequence 1243, Ap
33	33	91.7	7	10	US-09-989-789-1244	Sequence 1244, Ap
34	33	91.7	231	10	US-09-801-368-232	Sequence 232, App
35	32	88.9	7	9	US-09-990-186-996	Sequence 996, App
36	32	88.9	7	9	US-09-989-994-996	Sequence 996, App
37	32	88.9	7	10	US-09-989-789-996	Sequence 996, App
38	32	88.9	356	9	US-09-836-392-17	Sequence 17, Appl
39	31	86.1	7	9	US-10-006-069A-249	Sequence 249, App
40	31	86.1	7	9	US-09-846-033B-249	Sequence 249, App
41	31	86.1	7	9	US-09-990-186-1129	Sequence 1129, Ap
42	31	86.1	7	9	US-09-990-186-1130	Sequence 1130, Ap
43	31	86.1	7	9	US-09-990-186-1131	Sequence 1131, Ap
44	31	86.1	7	9	US-09-990-186-1132	Sequence 1132, Ap
45	31	86.1	7	9	US-09-990-186-1133	Sequence 1133, Ap

ALIGNMENTS

RESULT 1

US-10-006-069A-247
; Sequence 247, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-00583005
; CURRENT APPLICATION NUMBER: US/10/006.069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 247
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-247

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TSGHLSR 7

Db 1 TSGHLSR 7

RESULT 2

US-09-846-033B-247
; Sequence 247, Application US/09846033B
; Publication No. US2003004404A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE OF INVENTION: Finger Proteins
; FILE REFERENCE: 019496-005820US
; CURRENT APPLICATION NUMBER: US/09/846,033B
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 247
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-09-846-033B-247

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
Db 1 TSGHLSR 7

RESULT 3

US-09-990-186-1201
; Sequence 1201, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE OF INVENTION: TRIPLETS BY ZINC FINGERS
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1201
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example zfp
US-09-990-186-1201

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
Db 1 TSGHLSR 7

RESULT 4

US-09-989-994-1201
; Sequence 1201, Application US/09989994
; Publication No. US20030104526A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE OF INVENTION: TRIPLETS BY ZINC FINGERS
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,994
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1201
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example zfp
US-09-989-994-1201

Query Match 100.0%; Score 36; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
Db 1 TSGHLSR 7

RESULT 5

US-09-989-789-1201
; Sequence 1201, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE OF INVENTION: TRIPLETS BY ZINC FINGERS
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1201
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example zfp
US-09-989-789-1201

Query Match 100.0%; Score 36; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
Db 1 TSGHLSR 7

RESULT 6

US-10-006-069A-245
; Sequence 245, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc

; TITLE OF INVENTION: Finger Proteins
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006.069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 245
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-245

Query Match 91.7%; Score 33; DB 9; Length 7;
Best Local Similarity 85.7%; Pred. No. 4e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TSGHLSR 7
|||||
Db 1 TSGHLTR 7

RESULT 7
US-10-006-069A-248
; Sequence 248, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006.069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 248
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-10-006-069A-248

Query Match 91.7%; Score 33; DB 9; Length 7;
Best Local Similarity 85.7%; Pred. No. 4e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TSGHLSR 7
|||||
Db 1 TSGHLAR 7

RESULT 8

US-09-846-033B-245
; Sequence 245, Application US/09846033B
; Publication No. US2003004404A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005820US
; CURRENT APPLICATION NUMBER: US/09/846.033B
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 245
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-09-846-033B-245

Query Match 91.7%; Score 33; DB 9; Length 7;
Best Local Similarity 85.7%; Pred. No. 4e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TSGHLSR 7
|||||
Db 1 TSGHLTR 7

RESULT 9
US-09-846-033B-248
; Sequence 248, Application US/09846033B
; Publication No. US2003004404A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005820US
; CURRENT APPLICATION NUMBER: US/09/846.033B
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 248
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: finger
US-09-846-033B-248

Query Match 91.7%; Score 33; DB 9; Length 7;

Best Local Similarity 85.7%; Pred. No. 4e+05; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
Db 1 TSGHLR 7

RESULT 10

US-09-990-186-393
; Sequence 393, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:

; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 393
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

US-09-990-186-393
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP

Query Match 91.7%; Score 33; DB 9; Length 7;

Best Local Similarity 85.7%; Pred. No. 4e+05; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
Db 1 TSGHLR 7

RESULT 11

US-09-990-186-738
; Sequence 738, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:

; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 738
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

US-09-990-186-738
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP

Query Match 91.7%; Score 33; DB 9; Length 7;

Best Local Similarity 85.7%; Pred. No. 4e+05; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
Db 1 TSGHLR 7

RESULT 12

US-09-990-186-944
; Sequence 944, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:

; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 944
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

US-09-990-186-944
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP

Query Match 91.7%; Score 33; DB 9; Length 7;

Best Local Similarity 85.7%; Pred. No. 4e+05; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
Db 1 TSGHLR 7

RESULT 13

US-09-990-186-1036
; Sequence 1036, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:

; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1036
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

US-09-990-186-1036
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP

Query Match 91.7%; Score 33; DB 9; Length 7;

Best Local Similarity 85.7%; Pred. No. 4e+05; Mismatches 1; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
Db 1 TSGHLR 7

RESULT 14

US-09-990-186-1044
; Sequence 1044, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:

; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1044
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-990-186-1044

Query Match 91.7%; Score 33; DB 9; Length 7;
Best Local Similarity 85.7%; Pred. No. 4e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
|||||
Db 1 TSGHLTR 7

RESULT 15

US-09-990-186-1147
; Sequence 1147, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; TITLE OF INVENTION: TRIPLETS BY ZINC FINGERS
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1147
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-990-186-1147

Query Match 91.7%; Score 33; DB 9; Length 7;
Best Local Similarity 85.7%; Pred. No. 4e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TSGHLSR 7
|||||
Db 1 TSGHLTR 7

Search completed: July 15, 2003, 12:57:03
Job time : 16.3333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 16:30:06 ; Search time 69 seconds
(without alignments)
40.555 Million cell updates/sec

Title: SEQ55-247-68

Perfect score: 107

Sequence: 1 DRSNLRTSGHLRRSDHLR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	49.5	21	23	HIV-B zinc finger
2	53	49.5	532	22	Drosophila melanog
3	49	45.8	94	22	Propionibacterium
4	48.5	45.3	89	22	Propionibacterium
5	47	43.9	205	22	Propionibacterium
6	47	43.9	636	19	Aspergillus oryzae
7	47	43.9	636	19	Aspergillus oryzae
8	47	43.9	636	19	Aspergillus oryzae
9	46	43.0	228	22	5-aminotetralin a
10	45	42.1	1610	22	Human polypeptide
					C glutamicum prote

11	45	42.1	1610	22	AA76516	Corynebacterium gl
12	45	42.1	1610	22	AA76537	Corynebacterium gl
13	45	42.1	1856	20	AA21801	B. subtilis rib op
14	45	42.1	1856	21	AA83269	Polypeptide encode
15	44	41.1	21	20	AA33370	Zinc finger clone
16	44	41.1	52	22	AA46550	Propionibacterium
17	44	41.1	139	22	AA49761	Propionibacterium
18	44	41.1	1857	22	AA64454	Drosophila melanog
19	43	40.2	90	22	AA01814	Human polypeptide
20	43	40.2	105	22	AA36095	Klebsiella pneumonia
21	43	40.2	176	22	AA64891	Drosophila melanog
22	43	40.2	512	23	AB93563	Herbicidally activ
23	43	40.2	806	23	ABP27564	Streptococcus poly
24	42	39.3	21	20	AA33371	Zinc finger clone
25	42	39.3	21	23	AB05105	HIV-A zinc finger
26	42	39.3	21	23	AB05110	HIV-E zinc finger
27	42	39.3	63	23	AB09303	Human ORFX protein
28	42	39.3	71	22	AA55643	Propionibacterium
29	42	39.3	169	12	AA13403	Parvo virus B19 PA
30	42	39.3	331	21	AA36667	Arabidopsis thalia
31	42	39.3	351	21	AA36666	Arabidopsis thalia
32	42	39.3	373	21	AA36665	Arabidopsis thalia
33	42	39.3	444	22	AB68897	Drosophila melanog
34	42	39.3	680	23	AB92309	Herbicidally activ
35	42	39.3	701	21	AA39667	Arabidopsis thalia
36	42	39.3	705	21	AA39666	Arabidopsis thalia
37	42	39.3	742	21	AA39665	Arabidopsis thalia
38	42	39.3	1475	22	AA40232	Human polypeptide
39	42	39.3	1499	22	AA42018	Human polypeptide
40	41	38.3	140	22	ABG0239	Human ovarian anti
41	41	38.3	140	22	AA021861	Human cardiovascular
42	41	38.3	140	23	AA61710	Novel ovarian rela
43	41	38.3	183	22	AB17428	Human nervous syst
44	41	38.3	250	20	AA37675	Chlamydia trachoma
45	41	38.3	312	19	AAW5068	Streptococcus pneu

ALIGNMENTS

RESULT 1

ABB05107
ID ABB05107 standard; Peptide; 21 AA.

XX ABB05107;

XX 27-MAR-2002 (first entry)

DE HIV-B zinc finger amino acid sequence.

DE Human immunodeficiency virus; HIV; Herpesvirus; HSV; zinc finger;

KW nucleic acid binding protein; viral; promoter; infection; virucide;

KW anti-HIV.

OS Human immunodeficiency virus type 1.

OS Synthetic.

PN WO200185780-A2.

XX 15-NOV-2001.

XX 08-MAY-2001; 2001WO-GB02017.

XX 08-MAY-2000; 2000GB-0011068.

XX 30-MAY-2000; 2000GB-0013106.

XX 02-OCT-2000; 2000WO-GB03765.

XX 19-JAN-2001; 2001GB-0001446.

XX (GEND-) GENDAQ LTD.

XX Choo Y, Demaison C, Moore M, Papworth MA, Reynold L, Ullman CG;

PI Isalan M;

XX

pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA).

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct_sequences.

SQ Sequence 94 AA; Query Match 45.8%; Score 49; DB 22; Length 94; Best Local Similarity 64.7%; Pred. No. 1.5; Matches 11; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

Qy	6	TRTSG--HLSRRSDHLS	20
		:	:
Db	44	TRTTGHHQRRRADHLS	60

RESULT 4	-
AAU53591	-
ID	AAU53591 standard; Protein: 89 AA.

AAU53591;

27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #14487.

SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis; uveitis; endophthalmitis; bone; joint; central nervous system; ELISA; inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; dermatological; osteopathic; neuroprotectant.

OS *Propionibacterium acnes*:

PN WO200181581-A2.

PD 01-NOV-2001.

20-APR-2001: 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORIXA CORP.

PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A:

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

DR WPI; 2001-616774/71.

XX
XX

PT Propionibacterium acnes, polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.

PS Example 1; SEO ID No 14786: 1069pp; English.

Sequences AAU39105-AAU68017 represent *Propionibacterium* acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by *P. acnes*. The disorders include SAPHO syndrome (synovitis, acne).

pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA).

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct_sequences.

AA
SQ Sequence 89 AA;
Query Match 45.3%; Score 48.5; DB 22; Length 89;
Best Local Similarity 46.2%; Pred. No. 1.7;
Matches 12: Conservative 3: Mismatches 4: Indels 7: Gaps 1:

Qy 3 SNLTRTSGHLR-----RSDHLR 21
| : | : | | | | | : | | | |
pb 3 STVTVTGTSTRYRHGDCRRHLR 28

RESULT 5
AAU49653
ID AAU49653 standard: protein: 205 AA:

AA
AC: AAT49653:

DT 13-FEB-2002 (first entry)

XX
DE
Propionibacterium acnes immunogenic protein #10549.

SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis; kw ueitis; endophthalmitis; bone; joint; central nervous system; ELISA; kw inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; kw dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

AA WO200181581-A2.

01 - NOV - 2001 -

XX
PF 20-APR-2001: 2001WO-US12865

21-APR-2000: 2000US-199047P

PK 03-JUN-2000; 2000US-208841F
PR 07-JUL-2000; 2000US-216747P

PA (CORT-) CORTXA CORP.

XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A:

PI L'maisonneuve J, Zhang Y, Jen S, Carter D; Wang SK, Kiy IAW, Feisling DN, Mitterm UD, Wang SK

DR WP:I; 2001-616774/71.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
PT N FSDB, 88033549.

XX
PS
Example 1: SEQ ID No 10848: 1069pp: English.

Sequences AAU39105-AAU68017 represent *Propionibacterium* acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by *P. acnes*. The disorders include SAPHO syndrome (syndovitis, acne, AA

CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 205 AA;

Query Match 43.9%; Score 47; DB 22; Length 205;
 Best Local Similarity 44.8%; Pred. No. 7.8;
 Matches 13; Conservative 2; Mismatches 2; Indels 12; Gaps 2;

QY 1 DRSNLT-----RTSGHLSRRSDH 18
 : : : : :
 Db 93 DQHNVTVPQQRPDNDPRVSGH-SRRSDH 120

RESULT 6

AAW30557
 ID AAW30557 standard; Protein: 636 AA.

XX AC AAW30557;

XX DT 18-JAN-1999 (first entry)

XX DE Aspergillus oryzae 5-aminolevulinic acid synthase.

XX KW 5-Aminolevulinic acid synthase; hema gene; respiratory deficient;
 XX KW oxidative phosphorylation.

XX OS Aspergillus oryzae strain A1560 (IFO 4177).

XX FH Key Location/Qualifiers
 XX FT Peptide 1..35
 XX FT /note= "putative mitochondrial leader"

XX PN W09841640-A1.

XX PD 24-SEP-1998.

XX PF 17-MAR-1998; 98WO-US05156.

XX PR 17-MAR-1997; 97US-0819458.

XX PA (NOVO) NOVO NORDISK BIOTECH INC.
 XX PA (NOVO) NOVO-NORDISK AS.

XX PI Cherry JR, Elrod SL, Jensen EB;

XX DR WPI; 1998-521226/44.

XX DR N-PSDB; AAV45424.

XX PT Production of polypeptide in respiratory-deficient cells transformed
 XX PT with construct - that complements the defect and encodes
 XX PT polypeptide, also production of defective cells by disrupting gene
 XX PT essential for oxidative phosphorylation

XX PS Example 5; Fig 3A-B; 57pp; English.

XX CC This is the amino acid sequence of 5-aminolevulinic acid synthase
 XX CC (ALAS), an enzyme of the haem biosynthetic pathway, that can be
 XX CC isolated from Aspergillus oryzae strain A1560. The sequence was

CC deduced from an isolated genomic DNA (see AAV45424). A claimed
 CC method of producing a polypeptide comprises: (a) introducing into a
 CC respiratory-deficient cell mutant (i) one or more first nucleic
 CC acid sequences (NAS) which complement the respiratory defect and
 CC (ii) a second NAS which encodes the polypeptide, (b) cultivating
 CC the cell in a culture medium under aerobic conditions suitable for
 CC expression of the first and second NAS, and (c) isolating the
 CC polypeptide from the culture medium. The first NAS encodes a
 CC component of the electron transport chain, or an enzyme involved in
 CC the biosynthesis of ubiquinone, flavin or haem (e.g. ALAS). Also
 CC claimed is a respiratory-deficient mutant cell that is haem
 CC deficient (see AAV55426). The method is used to select and maintain
 CC transformed cells, to ensure high level expression and genetic
 CC stability of transformants during culture. It is applicable to all
 CC industrial fermentation processes, requiring only that the cells
 CC need oxygen for growth.

XX SQ Sequence 636 AA;

Query Match 43.9%; Score 47; DB 19; Length 636;
 Best Local Similarity 42.1%; Pred. No. 29;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 DRSNLTRTSGHLSRRSDHL 19
 : : : : :
 Db 531 ERLRITPTPGHIKEHRDHL 549

RESULT 7

AAW41509
 ID AAW41509 standard; Protein: 636 AA.

XX AC AAW41509;

XX DT 22-JUN-1998 (first entry)

XX DE Aspergillus oryzae 5-aminolevulinic acid synthase.

XX KW 5-Aminolevulinic acid synthase; hema gene; herbicide; haemoprotein.

XX OS Aspergillus oryzae strain A1560 (IFO 4177).

XX FH Key Location/Qualifiers
 XX FT Peptide 1..35
 XX FT /label= Sig_peptide
 XX FT /note= "mutative mitochondrial leader"

XX FT Peptide 11..15

XX FT /note= "haem regulatory motif"

XX FT Peptide 50..54

XX FT /note= "haem regulatory motif"

XX FT Binding-site 220..230
 XX FT /note= "glycine loop involved in pyridoxal
 XX FT phosphate cofactor binding"

XX PN W09747736-A1.

XX PD 18-DEC-1997.

XX PF 09-JUN-1997; 97WO-US09928.

XX PR 10-JUN-1996; 96US-0019399.

XX PA (NOVO) NOVO NORDISK BIOTECH INC.

XX PI Cherry JR, Elrod SL;

XX DR WPI; 1998-052303/05.

XX DR N-PSDB; AAV04102.

XX PT New 5-aminolaevulinic acid synthase from Aspergillus oryzae - used
 XX PT to produce 5-aminolaevulinic acid herbicide and to increase
 XX PT haemoprotein synthesis

PS Claim 1; Page 26-27; 62pp; English.

XX This protein comprises the 5-aminolevulinic acid synthase (ALA
CC synthase) of *Aspergillus niger* IFO 4177. Its amino acid sequence
CC was deduced from the isolate hema gene (see AAV04102) and shares
CC 81% identity with the *Aspergillus nidulans* ALA synthase. The
CC enzyme catalyses the reaction of glycine and succinyl-CoA to form
CC 5-aminolevulinic acid, which is useful as a herbicide. It is also
CC the rate-determining enzyme in biosynthesis of haem in liver cells
CC and differentiating erythrocytes, so overexpression in a cell can
CC be used to increase the yield of haemoproteins produced by the cell
CC in the absence of haem supplementation. Host cells, especially
CC bacterial, fungal, filamentous fungal and yeast cells, transformed
CC with a vector that includes the hema gene are claimed, and can be
CC used in a claimed method for producing ALA synthase.

XX Sequence 636 AA;

Query Match 43.9%; Score 47; DB 19; Length 636;

Best Local Similarity 42.1%; Pred. No. 29;

Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 DRSLNLTSGHLSRRSDHL 19

: | : | | |

Db 531 ERLRITTPGHIKEHRDHL 549

RESULT 8

AAW41498

ID AAW41498 standard; Protein: 636 AA.

AC AAW41498;

DT 11-JUN-1998 (first entry)

DE 5-aminolevulinic acid synthase.

XX 5-aminolevulinic acid synthase; porphobilinogen synthase;

KW haemoprotein production; filamentous fungus; haeme biosynthetic enzyme.

XX *Aspergillus oryzae*.

XX WO9747746-A1.

PN 18-DEC-1997.

XX 09-JUN-1997; 97WO-US10003.

XX 17-MAR-1997; 97US-0041158.

PR 10-JUN-1996; 96US-0662752.

XX (NOVO) NOVO NORDISK BIOTECH INC.

XX Cherry JR, Elrod SL, Jones A;

XX WPI; 1998-052312/05.

DR N-PSDB; AAV19709.

XX Method for the production of haemo:protein in a filamentous fungal
PT cell - by introducing a haeme biosynthetic enzyme coding sequence
PT and control sequences into the cell

XX Claim 9; Page 45-46; 113pp; English.

XX This sequence is the 5-aminolevulinic acid synthase of *Aspergillus*
CC *oryzae*. The DNA can be used in the method of the invention for
CC producing a haemoprotein. The method comprises: (a) introducing
CC into a filamentous fungal cell: (i) one or more control sequences capable
CC of directing the expression of a haeme biosynthetic enzyme encoded by a
CC nucleic acid sequence endogenous to the filamentous fungal cell, where
CC one or more of the control sequences are operably linked to the nucleic
CC acid sequence; and/or (ii) one or more copies of one or more second
CC nucleic acid sequences encoding a haeme biosynthetic enzyme: (b)

CC cultivating the filamentous fungal cell in a nutrient medium suitable for
CC production of the haemoprotein and the haeme biosynthetic enzymes; and
CC (c) recovering the haemoprotein from the nutrient medium of the
CC filamentous fungal cell. The method is used to yield commercially
CC significant quantities of haemoprotein in filamentous fungal strains.

XX Sequence 636 AA;

Query Match 43.9%; Score 47; DB 19; Length 636;

Best Local Similarity 42.1%; Pred. No. 29;

Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 DRSLNLTSGHLSRRSDHL 19

: | : | | |

Db 531 ERLRITTPGHIKEHRDHL 549

RESULT 9

AAW40688

ID AAW40688 standard; Protein: 228 AA.

XX AAW40688;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 5619.

XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX Zhao QA, Zhou P, Goodrich E, Drmanac RT;

XX WPI; 2001-442253/47.

DR N-PSDB; AAI59844.

XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -

XX Example 2; SEQ ID NO 5619; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAW38642-AAW42213) with nontropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX SQ Sequence 228 AA;
 Query Match 43.08; Score 46; DB 22; Length 228;
 Best Local Similarity 59.2%; Pred. NO. 13;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 NLRTSGHLSRRS 16
 DB 78 SLTASSGHLGRRS 90
 :||:|||||
 :||:|||||

RESULT 10
 AAG92912
 ID AAG92912 standard; Protein; 1610 AA.

XX AC AAG92912;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 6666.

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.

XX OS Corynebacterium glutamicum.

XX PN EP1108790-A2.

XX PD 20-JUN-2001.

XX PF 18-DEC-2000; 2000EP-0127688.

XX PR 16-DEC-1999; 99JP-0377484.

XX PR 07-APR-2000; 2000JP-0159162.

XX PR 03-AUG-2000; 2000JP-0280988.

XX PA (KYOW) KYOWA HAKKO KOGYO KK.

XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR WPI; 2001-376931/40.
 DR N-PSDB; AAH68131.

XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene

PS Claim 17; SEQ ID NO: 6666; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium glutamicum Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Corynebacterium, and identifying a homologue of a gene derived
 CC from corynebacterium. Corynebacterium bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.

XX SQ Sequence 1610 AA;

Query Match 42.18; Score 45; DB 22; Length 1610;
 Best Local Similarity 42.9%; Pred. NO. 1.8e+02;
 Matches 9; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 DRSNLRTSGHLSRRSDHLSR 21
 DB 1588 DEPEIGTVGAHLSRRIDEISR 1608
 :|||||:|
 :|||||:|

RESULT 11
 AAB76516

ID AAB76516 standard; Protein; 1610 AA.

XX AC AAB76516;

XX DT 11-APR-2001 (first entry)

XX Corynebacterium glutamicum MCT protein SEQ ID NO:14.

XX KW Corynebacterium glutamicum; brevivbacterium lactofermentum; MCT;
 KW membrane construction and membrane transport protein; petroleum spill;
 KW hydrocarbon degradation; gram positive aerobic bacterium; marker;
 KW identification; microorganism; fine chemical production; transformation;
 KW genome mapping; genetic engineering.

XX OS Corynebacterium glutamicum.

XX PN WO200100805-A2.

XX PD 04-JAN-2001.

XX PF 23-JUN-2000; 2000WO-IB00926.

XX PR 25-JUN-1999; 99US-0141031.

XX PR 08-JUL-1999; 99DE-1031454.

XX PR 08-JUL-1999; 99DE-1031478.

XX PR 08-JUL-1999; 99DE-1031563.

XX PR 09-JUL-1999; 99DE-1032122.

XX PR 09-JUL-1999; 99DE-1032124.

XX PR 09-JUL-1999; 99DE-1032125.

XX PR 09-JUL-1999; 99DE-1032128.

XX PR 09-JUL-1999; 99DE-1032180.

XX PR 09-JUL-1999; 99DE-1032182.

XX PR 09-JUL-1999; 99DE-1032190.

XX PR 09-JUL-1999; 99DE-1032191.

XX PR 09-JUL-1999; 99DE-1032209.

XX PR 09-JUL-1999; 99DE-1032212.

XX PR 09-JUL-1999; 99DE-1032227.

XX PR 09-JUL-1999; 99DE-1032228.

XX PR 09-JUL-1999; 99DE-1032229.

XX PR 09-JUL-1999; 99DE-1032230.

XX PR 14-JUL-1999; 99DE-1032927.

XX PR 14-JUL-1999; 99DE-1033005.

XX PR 14-JUL-1999; 99DE-1033006.

XX PR 27-AUG-1999; 99DE-1040764.

XX PR 27-AUG-1999; 99DE-1040765.

XX PR 27-AUG-1999; 99DE-1040766.

XX PR 27-AUG-1999; 99DE-1040830.

XX PR 27-AUG-1999; 99DE-1040831.

XX PR 27-AUG-1999; 99DE-1040832.

XX PR 27-AUG-1999; 99DE-1040833.

XX PR 31-AUG-1999; 99DE-1041378.

XX PR 31-AUG-1999; 99DE-1041379.

XX PR 31-AUG-1999; 99DE-1041395.

XX PR 03-SEP-1999; 99DE-1042077.

XX PR 03-SEP-1999; 99DE-1042078.

XX PR 03-SEP-1999; 99DE-1042079.

XX PR 03-SEP-1999; 99DE-1042088.

XX PA (BADI) BASF AG.

```

XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
PI
XX WPI; 2001-071486/08.
DR N-PSDB; AAF67749.
XX
XX Corynebacterium glutamicum nucleic acids encoding membrane construction
PT and membrane transport proteins or their portions, useful for typing or
PT identifying C. glutamicum or related bacteria, and as markers for
PT transformation.
XX
XX Claim 20; Page 153-158; 1119pp; English.
PS
XX AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
CC construction and membrane transport (MCT) proteins given in AAB76510 to
CC AAB76847. The MCT nucleic acids and proteins are useful in the
CC identification of microorganisms which can be used to produce fine
CC chemicals, for modulating fine chemical production in C. glutamicum or
CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or
CC identification of C. glutamicum or related bacteria, as reference points
CC for mapping C. glutamicum genome, and as markers for transformation.
CC AAF68082 and AAF68082 represent sequencing primers which are used in an
CC example from the present invention.
XX
XX Sequence 1610 AA;
PI
XX
XX Query Match 42.1%; Score 45; DB 22; Length 1610;
XX Best Local Similarity 42.9%; Pred. No. 1.8e+02;
XX Matches 9; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
QY 1 DRSNLTRTSGHLRRSDHLSR 21
DB 1588 DEPEIGTVGAHLRRIDEISR 1608
XX
XX RESULT 12
XX AAB76537
XX ID AAB76537 standard; Protein; 1610 AA.
XX
XX AC AAB76537;
XX
XX DT 11-APR-2001 (first entry)
XX
XX DE Corynebacterium glutamicum MCT protein SEQ ID NO:56.
XX
XX KW Corynebacterium glutamicum; Brevibacterium lactofermentum; MCT;
XX membrane construction and membrane transport protein; petroleum spill;
XX hydrocarbon degradation; gram positive aerobic bacterium; marker;
XX identification; microorganism; fine chemical production; transformation;
XX genome mapping; genetic engineering.
XX
XX OS Corynebacterium glutamicum.
XX
XX PN WO200100805-A2.
XX
XX PD 04-JAN-2001.
XX
XX PF 23-JUN-2000; 2000WO-IB00926.
XX
XX PR 25-JUN-1999; 99US-0141031.
XX PR 08-JUL-1999; 99DE-1031454.
XX PR 08-JUL-1999; 99DE-1031478.
XX PR 08-JUL-1999; 99DE-1031563.
XX PR 09-JUL-1999; 99DE-1032122.
XX PR 09-JUL-1999; 99DE-1032124.
XX PR 09-JUL-1999; 99DE-1032125.
XX PR 09-JUL-1999; 99DE-1032128.
XX PR 09-JUL-1999; 99DE-1032180.
XX PR 09-JUL-1999; 99DE-1032182.
XX PR 09-JUL-1999; 99DE-1032190.
XX PR 09-JUL-1999; 99DE-1032191.
XX PR 09-JUL-1999; 99DE-1032209.
XX PR 09-JUL-1999; 99DE-1032212.

```

```

PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032228.
PR 09-JUL-1999; 99DE-1032229.
PR 09-JUL-1999; 99DE-1032230.
PR 14-JUL-1999; 99DE-1032927.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040830.
PR 27-AUG-1999; 99DE-1040831.
PR 27-AUG-1999; 99DE-1040832.
PR 27-AUG-1999; 99DE-1040833.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041395.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042078.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042088.
XX
XX (BADI ) BASF AG.
XX
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
XX WPI; 2001-071486/08.
XX N-PSDB; AAF67770.
XX
XX Corynebacterium glutamicum nucleic acids encoding membrane construction
PT and membrane transport proteins or their portions, useful for typing or
PT identifying C. glutamicum or related bacteria, and as markers for
PT transformation.
XX
XX Claim 20; Page 253-258; 1119pp; English.
XX
XX AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
CC construction and membrane transport (MCT) proteins given in AAB76510 to
CC AAB76847. The MCT nucleic acids and proteins are useful in the
CC identification of microorganisms which can be used to produce fine
CC chemicals, for modulating fine chemical production in C. glutamicum or
CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or
CC identification of C. glutamicum or related bacteria, as reference points
CC for mapping C. glutamicum genome, and as markers for transformation.
CC AAF68082 and AAF68082 represent sequencing primers which are used in an
CC example from the present invention.
XX
XX Sequence 1610 AA;
XX
XX Query Match 42.1%; Score 45; DB 22; Length 1610;
XX Best Local Similarity 42.9%; Pred. No. 1.8e+02;
XX Matches 9; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
QY 1 DRSNLTRTSGHLRRSDHLSR 21
DB 1588 DEPEIGTVGAHLRRIDEISR 1608
XX
XX RESULT 13
XX AAY21801
XX ID AAY21801 standard; Protein; 1856 AA.
XX
XX AC AAY21801;
XX
XX DT 10-SEP-1999 (first entry)
XX
XX DE B. subtilis rib operon protein translated from reading frame 1.
XX Riboflavin; open reading frame; ORF; structural gene; promoter;
XX vitamin B2; Bacillus subtilis; rib operon.
XX
XX OS Bacillus subtilis.
XX

```


FT Misc-difference 721 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 948 /note= "Encoded by TAG stop codon"
FT
FT Misc-difference 938 /note= "Encoded by TAG stop codon"
FT
FT Misc-difference 965 /note= "Encoded by TAA stop codon"
FT
FT Misc-difference 974 /note= "Encoded by TAG stop codon"
FT
FT Misc-difference 1007 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1039 /note= "Encoded by TAA stop codon"
FT
FT Misc-difference 1041 /note= "Encoded by TAG stop codon"
FT
FT Misc-difference 1081 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1089 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1117 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1138 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1140 /note= "Encoded by TAA stop codon"
FT
FT Misc-difference 1146 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1153 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1181 /note= "Encoded by TAG stop codon"
FT
FT Misc-difference 1190 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1230 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1253 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1257 /note= "Encoded by TAA stop codon"
FT
FT Misc-difference 1270 /note= "Encoded by TAA stop codon"
FT
FT Misc-difference 1300 /note= "Encoded by TAG stop codon"
FT
FT Misc-difference 1303 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1342 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1345 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1376 /note= "Encoded by TAG stop codon"
FT
FT Misc-difference 1385 /note= "Encoded by TAA stop codon"
FT
FT Misc-difference 1386 /note= "Encoded by TAA stop codon"
FT
FT Misc-difference 1409 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1411 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1420 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1435 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1457 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1481 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1507 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1516 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1520 /note= "Encoded by TGA stop codon"
FT

FT Misc-difference 1531 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1537 /note= "Encoded by TAA stop codon"
FT
FT Misc-difference 1628 /note= "Encoded by TAA stop codon"
FT
FT Misc-difference 1641 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1645 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1659 /note= "Encoded by TAA stop codon"
FT
FT Misc-difference 1684 /note= "Encoded by TAA stop codon"
FT
FT Misc-difference 1776 /note= "Encoded by TGA stop codon"
FT
FT Misc-difference 1808 /note= "Encoded by TAA stop codon"
FT
FT Misc-difference 1808 /note= "Encoded by TAA stop codon"
XX
XX EP1001026-A2.
PN
XX 17-MAY-2000.
PD
XX
XX 22-JUN-1990; 99EP-0124888.
PF
XX
XX 22-JUN-1989; 89US-0370378.
PR
XX 22-JUN-1990; 90EP-0111916.
PR
XX
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA
XX Perkins JB, Pero JG, Sloma A;
PI
Query Match 42.1%; Score 45; DB 21; Length 1856;
Best Local Similarity 55.0%; Pred. No. 2.2e+02;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY 1 DRNLRTTSGHLSRRSDHLS 20
DB 385 DRPNPIRSSALLSXRTDKLS 404
RESULT 15
AAY33370
ID AAY33370 standard; peptide; 21 AA.
XX
AC AAY33370;
XX
XX 01-DEC-1999 (first entry)
XX
XX Zinc finger clone zfhHA(Y) peptide.
XX
XX Zinc finger; DNA binding; Cys2-His2 class; 5-methylcytosine; meC;
KW diagnostic; detection; chimera.
XX
XX Unidentified.
XX
XX WO9947656-A2.
XX
XX 23-SEP-1999.
XX
XX 17-MAR-1999; 99WO-GB00816.
XX
XX 17-MAR-1998; 98GB-0005576.
PR
XX 31-MAR-1998; 98GB-0006895.
PR
XX 03-APR-1998; 98GB-0007246.
XX
XX (MEDI-) MEDICAL RES COUNCIL.
XX
XX Choo Y, Isalan M;
PI
XX WPI; 1999-562106/47.
XX
XX

PT New zinc finger polypeptides that bind DNA containing modified bases,
PT used as diagnostic and research reagents and for regulating gene
PT transcription
XX
PS Example 4; Page 36; 56pp; English.
XX
CC This invention describes a novel zinc finger (ZF) polypeptide (I) that
CC binds to a target DNA sequence (II) containing a modified base but not to
CC an otherwise identical sequence containing the equivalent unmodified
CC base. The invention also describes methods for preparing a DNA-binding
CC polypeptide of the Cys2-His2 ZF class, able to recognize sequences
CC containing a 5-methylcytosine (mec) residue. (I) are used as diagnostic
CC reagents (for detecting modified nucleic acids in complex mixtures,
CC including differentiation of single-base modifications), in research and
CC to produce chimeras, e.g. by fusion to a catalytic domain of a
CC restriction enzyme (the product can then cleave only modified DNA), or to
CC a DNA cleavage or activating domain (to give products that can regulate
CC gene transcription, by sequence-specific cleavage or activation,
CC dependent on presence of a modified base). (I) recognize modified bases
CC in preference to unmodified ones, in a sequence-dependent manner, so have
CC extremely high specificity. This sequence represents a zinc finger
CC clone peptide fragment described in the method of the invention.

XX Sequence 21 AA;

Query Match 41.1%; Score 44; DB 20; Length 21;
Best Local Similarity 75.0%; Pred. No. 1.8;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 9 SGHLRRSRSDHLS 20
| | : |||||
Db 2 SDELTRRSRDHLS 13

Search completed: July 15, 2003, 16:32:41
Job time : 70 secs.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 16:30:06 : Search time 15 Seconds
(without alignments)
134.588 Million cell updates/sec

Title: SEQ55-247-68

Perfect score: 107

Sequence: 1 DRSNLRTSGHLRRSDHLSR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	50.5	87	F45714	probable regulator
2	54	50.5	87	D61547	hypothetical prote
3	54	50.5	240	E45714	probable regulator
4	54	50.5	241	C46181	px-tax-orf II (alt
5	52	48.6	328	T15555	G protein-coupled
6	46	43.0	71	D37390	transfer protein T
7	45	42.1	75	BVECRY	tray protein - Esc
8	45	42.1	502	E81780	L-aspartate oxidas
9	45	42.1	692	A01857	hypothetical prote
10	44.5	41.6	557	T07632	isocitrate lyase (
11	44.5	41.6	558	T07631	isocitrate lyase (
12	44.5	41.6	575	T06353	isocitrate lyase (
13	44.5	41.6	576	1 WZCNU	isocitrate lyase (
14	44.5	41.6	576	1 WZCSI	isocitrate lyase (
15	44.5	41.6	576	1 WZPPI	isocitrate lyase (
16	44.5	41.6	576	1 S53505	isocitrate lyase -
17	44	41.1	542	A84229	sensory histidine
18	43	40.2	136	D35845	transformer-2 sex-
19	43	40.2	256	A32373	transformer-2 sex-
20	43	40.2	264	A31638	transformer-2 sex-
21	43	40.2	462	1 C71296	glycine-tRNA ligas
22	43	40.2	1122	2 S25563	DNA-directed RNA p
23	42	39.3	213	2 A84248	transcription regu
24	42	39.3	226	2 H71700	hypothetical prote
25	42	39.3	425	2 B83412	hypothetical prote
26	42	39.3	462	2 S55114	hypothetical prote
27	42	39.3	502	2 H81203	L-aspartate oxidas
28	42	39.3	545	2 A03330	L-aspartate oxidas
29	42	39.3	608	2 F84430	hypothetical prote

ALIGNMENTS

RESULT 1

F45714

probable regulatory protein pl3 II, alternative splice form - human T-cell lymphotrop
C:Species: human T-cell lymphotropic virus type 1, HTLV-1
C:Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 23-May-1997
C:Accession: F45714

R:Gessain, A.; Boeri, E.; Yanagihara, R.; Gallo, R.C.; Franchini, G.

J. Virol. 67, 1015-1023, 1993

A:Title: Complete nucleotide sequence of a highly divergent human T-cell leukemia (ly
from other geographical regions.

A:Reference number: A45714; MUID:93124536; PMID:8419636

A:Contents: HTLV-IMEL5

A:Accession: F45714

A:Status: Preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-87 <GES>

A:Note: sequence extracted from NCBI backbone (NCBIP:122473)

Query Match 50.5%; Score 54; DB 2; Length 87;

Best Local Similarity 71.4%; Pred. No. 0.13; Mismatches 2; Indels 0; Gaps 0;

Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 TRTSGHLRRSDHL 19

:| :||||| ||||

Db 46 SRPTGHLRRSDHL 59

RESULT 2

D61547

hypothetical protein II (px region) - human T-cell lymphotropic virus type 1 (isolate
C:Species: human T-cell lymphotropic virus type 1, HTLV-1
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 08-Oct-1999

C:Accession: D61547

R:Tsujiimoto, A.; Teruuchi, T.; Imamura, J.; Shimotohno, K.; Miyoshi, I.; Miwa, M.

Mol. Biol. Med. 5, 29-42, 1988

A:Title: Nucleotide sequence analysis of a provirus derived from HTLV-1-associated my

A:Reference number: A61547; MUID:88232270; PMID:2897612

A:Accession: D61547

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-87 <TSU>

A:Cross-references: GB:M37301; NID:g541634; PIDN:AAA45391.1; PID:g541638

Query Match

Best Local Similarity 50.5%; Score 54; DB 2; Length 87;

Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 TRTSGHLRRSDHL 19

:| :||||| ||||

Db 46 SRPTGHLRRSDHL 59

RESULT 3

E45714
probable regulatory function protein p30 II, alternative splice form - human T-cell lymph
C:Species: human T-cell lymphotropic virus type 1, HTLV-1
C:Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 23-May-1997
C:Accession: E45714
R:Gessain, A.; Boeri, E.; Yanggihara, R.; Gallo, R.C.; Franchini, G.
J. Virol. 67, 1015-1023, 1993
A:Title: Complete nucleotide sequence of a highly divergent human T-cell leukemia (lymph
from other geographical regions.
A:Reference number: A45714; MUID:93124536; PMID:8419636
A:Contents: HTLV-IMEL5
A:Accession: E45714
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-240 <GES>
A:Note: sequence extracted from NCBI backbone (NCBIP:122472)

Query Match 50.5%; Score 54; DB 2; Length 240;
Best Local Similarity 71.4%; Pred. No. 0.4;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 TRTSGHLRSRSDHL 19
: : ||||| ||||
Db 199 SRPTGHLRSRSDHL 212

RESULT 4

E46181
px-tax-orf II (alternatively spliced) - human T-cell lymphotropic virus type 1
C:Species: human T-cell lymphotropic virus type 1, HTLV-1
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 17-Nov-2000
C:Accession: E46181
R:Koralnik, I.J.; Gessain, A.; Klotman, M.E.; Lo Monaco, A.; Berneman, Z.N.; Franchini,
Proc. Natl. Acad. Sci. U.S.A. 89, 8813-8817, 1992
A:Title: Protein isoforms encoded by the px region of human T-cell leukemia/lymphotropic
A:Reference number: A46181; MUID:92409607; PMID:1528897
A:Accession: E46181
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-241 <KOR>
A:Note: sequence extracted from NCBI backbone (NCBIP:114307)
A:Accession: E46181
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 155-241 <KO2>
A:Note: sequence extracted from NCBI backbone (NCBIP:114308)

Query Match 50.5%; Score 54; DB 2; Length 241;
Best Local Similarity 71.4%; Pred. No. 0.4;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 TRTSGHLRSRSDHL 19
: : ||||| ||||
Db 200 SRPTGHLRSRSDHL 213

RESULT 5

T15555
G protein-coupled receptor homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15555
R:Latrille, P.
submitted to the EMBL Data Library, May 1996
A:Description: The sequence of C. elegans cosmid C18F10.
A:Reference number: Z18369
A:Accession: T15555
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-328 <LAT>
A:Cross-references: EMBL:U00049; NID:g485101; PID:g485105; PIDN:AAC47053.1; GSPDB:GN0002
A:Experimental source: strain Bristol N2; clone C18F10

C:Genetics:

A:Gene: CESP:src-1
A:Map position: 3
A:Introns: 69/3; 190/3; 298/3

Query Match 48.6%; Score 52; DB 2; Length 328;
Best Local Similarity 68.8%; Pred. No. 1.2;
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 NLRTSGHLRSRSDHL 19
: : ||||| ||||
Db 283 NFRTSGHLGYSKHL 298

RESULT 6

D37390
transfer protein Tray - Escherichia coli plasmid P307
C:Species: Escherichia coli
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-Aug-1999
C:Accession: D37390
R:Gaus-Goeldner, A.; Gaus, H.; Schlacher, T.; Hoegenauer, G.
Plasmid 24, 119-131, 1990
A:Title: The sequences of genes bordering orit in the enterotoxin plasmid P307: compa
A:Reference number: A37390; MUID:91261994; PMID:2096398
A:Accession: D37390
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-71 <GRA>
A:Cross-references: GB:M62986; NID:g150463; PIDN:AAA25524.1; PID:g150467
C:Genetics:
A:Genome: plasmid
A:Superfamily: tray protein
C:Keywords: DNA binding

Query Match 43.0%; Score 46; DB 2; Length 71;
Best Local Similarity 47.6%; Pred. No. 2;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 DRSLNTRTSGHLRSRSDHL 21
||| ||: | |||:
Db 31 DRSGRTKNEVFIRLADHLNR 51

RESULT 7

BVPCRY
tray protein - Escherichia coli plasmids
C:Species: Escherichia coli
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: C25033; C32014
R:Finlay, B.B.; Frost, L.S.; Paranchych, W.
J. Bacteriol. 168, 132-139, 1986
A:Title: Origin of transfer of Incf plasmids and nucleotide sequences of the type II
A:Reference number: A25033; MUID:87008371; PMID:3531163
A:Accession: C25033
A:Molecule type: DNA
A:Residues: 1-75 <FIN>
A:Cross-references: GB:M15136; NID:g151788; PIDN:AAA26076.1; PID:g151789
A:Experimental source: plasmid R100-1
R:Inamoto, S.; Yoshioke, Y.; Ohtsubo, E.
J. Bacteriol. 170, 2749-2757, 1988
A:Title: Identification and characterization of the products from the tray and tray g
A:Reference number: A32014; MUID:88227859; PMID:2836369
A:Accession: C32014
A:Molecule type: DNA
A:Residues: 1-75 <INA>
A:Cross-references: GB:M20941; NID:g151778; PIDN:AAA26073.1; PID:g151781
A:Experimental source: plasmid R100
C:Genetics:
A:Gene: tray
A:Genome: plasmid
A:Start codon: TTG
C:Function:
A:Description: Involved in the conjugation process of bacterial cells for the exchang


```

QY      1  DRNLRTSGHLSRR-----SDHLSR 21
      ||::|:  ||::  |::|:|
Db      205  DQSSVTKKCGHMAGKVLVAISEHINR 230

RESULT 15
WZRP1
Isocitrate lyase (EC 4.1.3.1) - rape
N:Alternate names: isocitrase; isocitratase; isocitritase
C:Species: Brassica napus (rape)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 30-Jun-1993
C:Accession: JQ1105
R:Comai, L.; Dietrich, R.A.; Maslyar, D.J.; Baden, C.S.; Harada, J.J.
  Plant Cell 1, 293-300, 1989
A:Title: Coordinate expression of transcriptionally regulated isocitrate lyase and ma
A:Reference number: JQ1105; MUID:92393389; PMID:2535504
A:Accession: JQ1105
A:Molecule type: mRNA
A:Residues: 1-576 <COM>
C:Comment: This enzyme catalyzes the reversible conversion of isocitrate to succinate
  is involved in storage lipid mobilization during the growth of higher plant seedling
C:Superfamily: isocitrate lyase
C:Keywords: carbon-carbon lyase; glyoxylate bypass; oxo-acid-lyase

Query Match      41.6%; Score 44.5; DB 1; Length 576;
Best Local Similarity 30.8%; Pred. No. 34;
Matches      8; Conservative 10; Mismatches 3; Indels 5; Gaps 1;

QY      1  DRNLRTSGHLSRR-----SDHLSR 21
      ||::|:  ||::  |::|:|
Db      205  DQSSVTKKCGHMAGKVLVAISEHINR 230

Search completed: July 15, 2003, 16:31:25
Job time : 17 secs

```



```
ID TRY5_ECOLI STANDARD; PRT; 71 AA.
AC Q00736;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Tray protein.
GN TRAY.
OS Escherichia coli.
OG Plasmid IncFII R100.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=711;
RX MEDLINE=91261994; PubMed=2096398;
RA Graus-Goeldner A., Graus H., Schlacher T., Hoegenauer G.;
RT "The sequences of genes bordering orit in the enterotoxin plasmid
PL307: comparison with the sequences of plasmids F and RL.";
RL Plasmid 24:119-131(1990).
CC -!- FUNCTION: INVOLVED IN THE CONJUGATION PROCESS OF BACTERIAL CELLS
CC FOR THE EXCHANGE OF PLASMID DNA. IT IS ALSO RESPONSIBLE FOR
CC CONJUGAL DNA METABOLISM. TRAY IS REQUIRED FOR STRAND-SPECIFIC
CC NICKING AT ORIT, THE TRANSFER ORIGIN.
CC -!- SIMILARITY: TO TRAY PROTEIN OF OTHER PLASMIDS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M62986; AAA25524.1; -
CC PIR; D37390; D37390.
CC Plasmid; Conjugation; DNA-binding.
CC SEQUENCE 71 AA; 8103 MW; 9A031E0ECBE99263 CRC64;
Query Match 43.08; Score 46; DB 1; Length 71;
Best Local Similarity 47.68; Pred. No. 0.48;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 1 DRSNLTGSLRRSDHLSR 21
III I:I: I III:I
DB 31 DRSGRKTNEVFLKDLNLR 51
RESULT 3
ID TRY3_ECOLI STANDARD; PRT; 75 AA.
AC P05835;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Tray protein.
GN TRAY.
OS Escherichia coli.
OG Plasmid IncFII R100-1, and Plasmid IncFII R100.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC PLASMID=IncFII R100-1;
RX MEDLINE=87008371; PubMed=3531163;
RA Finlay B.B., Frost L.S., Paranchych W.;
RT "Origin of transfer of IncF plasmids and nucleotide sequences of the
type II orit, traM, and traY alleles from ColB4-K98 and the type IV
tray allele from R100-1.";
RL J. Bacteriol. 168:132-139(1996).
RN [2]
RP SEQUENCE FROM N.A.
```

```
RC PLASMID=IncFII R100;
RX MEDLINE=88227859; PubMed=2836369;
RA Inamoto S., Yoshioka Y., Ohtsubo E.;
RT "Identification and characterization of the products from the traJ
and traY genes of plasmid R100.";
RL J. Bacteriol. 170:2749-2757(1988).
CC -!- FUNCTION: INVOLVED IN THE CONJUGATION PROCESS OF BACTERIAL CELLS
CC FOR THE EXCHANGE OF PLASMID DNA. IT IS ALSO RESPONSIBLE FOR
CC CONJUGAL DNA METABOLISM. TRAY IS REQUIRED FOR STRAND-SPECIFIC
CC NICKING AT ORIT, THE TRANSFER ORIGIN.
CC -!- SIMILARITY: TO TRAY PROTEIN OF OTHER PLASMIDS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M15136; AAA26076.1; -
CC PIR; M20941; AAA26073.1; -
CC PIR; C25033; BVECRY.
CC PIR; C32014; C32014.
CC Plasmid; Conjugation; DNA-binding.
CC SEQUENCE 75 AA; 8542 MW; 88D4B04C4B5DE07A CRC64;
Query Match 42.18; Score 45; DB 1; Length 75;
Best Local Similarity 47.68; Pred. No. 0.74;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 1 DRSNLTGSLRRSDHLSR 21
III I:I: I III:I
DB 31 ERSGRYKTNVFLRLDLNLR 51
RESULT 4
ID ACE2_SOYBN STANDARD; PRT; 557 AA.
AC P45437;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Isocitrate lyase 2 (EC 4.1.3.1) (Isocitrate 2) (Isocitratease 2)
DE (ICL 2) (Fragment).
DE ICL2.
GN Glycine max (Soybean).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Maple Arrow; TISSUE=Cotyledon;
RA Guex N., Henry H., Widmer F.;
RL Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN STORAGE LIPID MOBILIZATION DURING THE GROWTH
CC OF HIGHER PLANT SEEDLING (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Isocitrate -> succinate + glyoxylate.
CC -!- PATHWAY: FIRST STEP IN GLYOXYLATE BYPASS, AN ALTERNATIVE TO THE
CC TRICARBOXYLIC ACID CYCLE (IN BACTERIA, PLANTS, AND FUNGI).
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: GLYOXYSOMAL.
CC -!- SIMILARITY: BELONGS TO THE ISOCITRATE LYASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```

DR EMBL; L02330; AAA33977.1; -
DR HSSP; P28298; IDQU.
DR InterPro; IPR000918; Isocit_lyase.
DR Pfam; PF00463; ICL; 1.
DR ProDom; PD001857; Isocit_lyase; 2.
DR PROSITE; PS00161; ISOCITRATE_LYASE; 1.
DR PROSITE; PS00342; MICROBODIES_CTER; 1.
KW Glyoxylate bypass; Tricarboxylic acid cycle; Lyase; Glyoxysome;
KW Multigene family.
FT NON_TER 1
FT ACT_SITE 195 195 BY SIMILARITY.
FT SITE 555 557 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 557 AA; 62840 MW; 486D5CE420B7AD95 CRC64;

Query Match 41.6%; Score 44.5; DB 1; Length 557;
Best Local Similarity 30.8%; Pred. No. 9.3;
Matches 8; Conservative 10; Mismatches 3; Indels 5; Gaps 1;

QY 1 DRSNLTSTSGHLSRR-----SDHLSR 21
|:|:|: |:|:|: |:|:|:
Db 187 DQSSVTRKCGHMGKVLVAISEHINR 212

RESULT 5
ACEL_SOYBN STANDARD; PRT; 558 AA.
ID ACCEL_SOYBN
AC P45456;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Isocitrate lyase 1 (EC 4.1.3.1) (Isocitrate 1) (Isocitrate 1)
DE (ICL 1) (fragment).
GN ICL1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Maple Arrow; TISSU= Cotyledon;
RA Guex N., Henry H., Widmer F.;
RL Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN STORAGE LIPID MOBILIZATION DURING THE GROWTH
CC OF HIGHER PLANT SEEDLING (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Isocitrate -> succinate + glyoxylate.
CC -!- PATHWAY: FIRST STEP IN GLYOXYLATE BYPASS, AN ALTERNATIVE TO THE
CC TRICARBOXYLIC ACID CYCLE (IN BACTERIA, PLANTS, AND FUNGI).
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Glyoxysomal.
CC -!- SIMILARITY: BELONGS TO THE ISOCITRATE LYASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L02329; AAA33976.1; -
DR HSSP; P28298; IDQU.
DR InterPro; IPR000918; Isocit_lyase.
DR Pfam; PF00463; ICL; 1.
DR ProDom; PD001857; Isocit_lyase; 2.
DR PROSITE; PS00161; ISOCITRATE_LYASE; 1.
DR PROSITE; PS00342; MICROBODIES_CTER; 1.
KW Glyoxylate bypass; Tricarboxylic acid cycle; Lyase; Glyoxysome;
KW Multigene family.
FT NON_TER 1
FT ACT_SITE 195 195 BY SIMILARITY.
FT SITE 556 558 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 558 AA; 62845 MW; 99B5CA5EA24F5DDB CRC64;

```

```

Query Match 41.6%; Score 44.5; DB 1; Length 558;
Best Local Similarity 30.8%; Pred. No. 9.3;
Matches 8; Conservative 10; Mismatches 3; Indels 5; Gaps 1;

QY 1 DRSNLTSTSGHLSRR-----SDHLSR 21
|:|:|: |:|:|: |:|:|:
Db 187 DQSSVTRKCGHMGKVLVAISEHINR 212

RESULT 6
ACEA_DENCR STANDARD; PRT; 574 AA.
ID ACEA_DENCR
AC Q9SE26;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Isocitrate lyase (EC 4.1.3.1) (Isocitrate) (ICL).
GN ICL.
OS Dendrobium crumenatum (Tropical pigeon orchid).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
OC Epidendroideae; higher Epidendroideae; Dendrobleae; Dendroblinae;
OC Dendrobium.
OX NCBI_TaxID=51096;
RN [1]
RP SEQUENCE FROM N.A.
RA Vellupillai M., Goh C.-J., Swarup S.;
RT "Sequence analysis of Dcricl, an isocitrate lyase gene from the
RT tropical orchid, Dendrobium crumenatum."
RL (In) Plant Gene Register PGF99-178.
CC -!- CATALYTIC ACTIVITY: Isocitrate -> succinate + glyoxylate.
CC -!- PATHWAY: FIRST STEP IN GLYOXYLATE BYPASS, AN ALTERNATIVE TO THE
CC TRICARBOXYLIC ACID CYCLE (IN BACTERIA, PLANTS, AND FUNGI).
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Glyoxysomal.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED MAXIMALLY DURING POSTGERMINATIVE
CC GROWTH.
CC -!- SIMILARITY: BELONGS TO THE ISOCITRATE LYASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF193815; AAF04598.1; -
DR HSSP; P28298; IDQU.
DR InterPro; IPR000918; Isocit_lyase.
DR Pfam; PF00463; ICL; 1.
DR ProDom; PD001857; Isocit_lyase; 2.
DR PROSITE; PS00161; ISOCITRATE_LYASE; 1.
DR PROSITE; PS00342; MICROBODIES_CTER; 1.
KW Glyoxylate bypass; Tricarboxylic acid cycle; Lyase; Glyoxysome.
FT ACT_SITE 212 212 PROBABLE.
FT SITE 572 574 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 574 AA; 63993 MW; 0B568ADE3248DE2E CRC64;

Query Match 41.6%; Score 44.5; DB 1; Length 574;
Best Local Similarity 30.8%; Pred. No. 9.6;
Matches 8; Conservative 10; Mismatches 3; Indels 5; Gaps 1;

QY 1 DRSNLTSTSGHLSRR-----SDHLSR 21
|:|:|: |:|:|: |:|:|:
Db 204 DQSSVTRKCGHMGKVLVAISEHINR 229

RESULT 7
ACEA_LYCES STANDARD; PRT; 575 AA.
ID ACEA_LYCES
AC P49297;

```



```

RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE-93241152; PubMed-8479425;
RA  Zhang J.Z., Genez-Pedrozo M., Baden C.S., Harada J.J.;
RT  "Two classes of isocitrate lyase genes are expressed during late
RL  embryogeny and postgermination in Brassica napus L.";
RN  Mol. Gen. Genet. 238:177-184(1993).
RN  [3]
RP  SEQUENCE FROM N.A.
RX  STRAIN-CV. Global; TISSUE-Cotyledon;
RA  Olesen C., Thomsen K.K., Svendsen I., Brandt A.;
RL  Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC  -1- FUNCTION: INVOLVED IN STORAGE LIPID MOBILIZATION DURING THE GROWTH
CC  OF HIGHER PLANT SEEDLING.
CC  -1- CATALYTIC ACTIVITY: Isocitrate - succinate + glyoxylate.
CC  -1- PATHWAY: FIRST STEP IN GLYOXYLATE BYPASS, AN ALTERNATIVE TO THE
CC  TRICARBOXYLIC ACID CYCLE (IN BACTERIA, PLANTS, AND FUNGI).
CC  -1- SUBUNIT: HOMOTETRAMER.
CC  -1- SIMILARITY: BELONGS TO THE ISOCITRATE LYASE FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  EMBL; L08482; AAA32992.1; -
CC  EMBL; Y13356; CAA73792.1; -
CC  PIR; JQ1105; WZRPI.
CC  HSSP; P28298; IDQU.
CC  InterPro: IPR000918; Isocit_lyase.
CC  Pfam; PF00463; ICL; 1.
CC  ProDom; PD001857; Isocit_lyase; 2.
CC  PROSITE; PS00161; ISOCITRATE_LYASE; 1.
CC  PROSITE; PS00342; MICROBODIES_CTER; 1.
CC  Glyoxylate bypass; Tricarboxylic acid cycle; Lyase; Glyoxysome.
CC  ACT_SITE 213 213 PROBABLE.
CC  SITE 574 576 MICROBODY TARGETING SIGNAL (POTENTIAL).
CC  FT CONFLICT 75 75 A -> V (IN REF. 3).
CC  FT CONFLICT 177 177 D -> G (IN REF. 3).
CC  FT CONFLICT 256 256 A -> P (IN REF. 3).
CC  FT CONFLICT 278 278 S -> N (IN REF. 3).
CC  FT CONFLICT 340 340 N -> T (IN REF. 3).
CC  FT CONFLICT 344 344 N -> I (IN REF. 3).
CC  SQ SEQUENCE 576 AA; 64325 MW; 81A87701A4ACC350 CRC64;

Query Match 41.6%; Score 44.5; DB 1; Length 576;
Best Local Similarity 30.8%; Pred. No. 9.7;
Matches 8; Conservative 10; Mismatches 3; Indels 5; Gaps 1;

QY 1 DRSNLRTSGHLSRR-----SDHLSR 21
Db 205 DQSSVTKKCGHMAGKVLVAVSEHNR 230

RESULT 10
ACEA_CUCMA STANDARD; PRT; 576 AA.
ID ACEA_CUCMA
AC P31110;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Isocitrate lyase (EC 4.1.3.1) (isocitrate)
OS Cucurbita maxima (Pumpkin) (Winter squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3661;

```

```

RN  [1]
RP  SEQUENCE FROM N.A.
RX  TISSUE-Cotyledon;
RA  MEDLINE-97134001; PubMed-8979395;
RA  Mano S., Hayashi M., Kondo M., Nishimura M.;
RT  "cDNA cloning and expression of a gene for isocitrate lyase in pumpkin
RL  cotyledons.";
RN  Plant Cell Physiol. 37:941-948(1996).
RN  CC -1- FUNCTION: INVOLVED IN STORAGE LIPID MOBILIZATION DURING THE GROWTH
RN  OF HIGHER PLANT SEEDLING.
RN  CC -1- CATALYTIC ACTIVITY: Isocitrate - succinate + glyoxylate.
RN  CC -1- PATHWAY: FIRST STEP IN GLYOXYLATE BYPASS, AN ALTERNATIVE TO THE
RN  TRICARBOXYLIC ACID CYCLE (IN BACTERIA, PLANTS, AND FUNGI).
RN  CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
RN  CC -1- SUBCELLULAR LOCATION: Glyoxysomal.
RN  CC -1- SIMILARITY: BELONGS TO THE ISOCITRATE LYASE FAMILY.
RN  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RN  between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN  the European Bioinformatics Institute. There are no restrictions on its
RN  use by non-profit institutions as long as its content is in no way
RN  modified and this statement is not removed. Usage by and for commercial
RN  entities requires a license agreement (See http://www.isb-sib.ch/announce/
RN  or send an email to license@isb-sib.ch).
RN  EMBL; D78256; BAA11320.1; -
RN  HSSP; P28298; IDQU.
RN  InterPro: IPR000918; Isocit_lyase.
RN  Pfam; PF00463; ICL; 1.
RN  ProDom; PD001857; Isocit_lyase; 2.
RN  PROSITE; PS00161; ISOCITRATE_LYASE; 1.
RN  PROSITE; PS00342; MICROBODIES_CTER; 1.
RN  Glyoxylate bypass; Tricarboxylic acid cycle; Lyase; Glyoxysome.
RN  ACT_SITE 213 213 BY SIMILARITY.
RN  SITE 574 576 MICROBODY TARGETING SIGNAL (POTENTIAL).
RN  SQ SEQUENCE 576 AA; 64358 MW; 269D93D4F13B5990 CRC64;

Query Match 41.6%; Score 44.5; DB 1; Length 576;
Best Local Similarity 30.8%; Pred. No. 9.7;
Matches 8; Conservative 10; Mismatches 3; Indels 5; Gaps 1;

QY 1 DRSNLRTSGHLSRR-----SDHLSR 21
Db 205 DQSSVTKKCGHMAGKVLVAVSEHNR 230

RESULT 11
ACEA_CUCSA STANDARD; PRT; 576 AA.
ID ACEA_CUCSA
AC F49296;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Isocitrate lyase (EC 4.1.3.1) (isocitrate) (Isocitratase) (ICL).
OS Cucumis sativus (Cucumber).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Cucurbitales; Cucurbitaceae; Cucumis.
OX NCBI_TaxID=3659;

RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN-CV. Masterpiece; TISSUE=Leaf;
RA  MEDLINE-95201243; PubMed-7894014;
RA  Reynolds S.J., Smith S.M.;
RT  "The isocitrate lyase gene of cucumber: isolation, characterisation
RL  and expression in cotyledons following seed germination.";
RN  Plant Mol. Biol. 27:487-497(1995).
RN  CC -1- FUNCTION: INVOLVED IN STORAGE LIPID MOBILIZATION DURING THE GROWTH
RN  OF HIGHER PLANT SEEDLING.
RN  CC -1- CATALYTIC ACTIVITY: Isocitrate - succinate + glyoxylate.
RN  CC -1- PATHWAY: FIRST STEP IN GLYOXYLATE BYPASS, AN ALTERNATIVE TO THE
RN  TRICARBOXYLIC ACID CYCLE (IN BACTERIA, PLANTS, AND FUNGI).
RN  CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).

```

```
CC -1- SUBCELLULAR LOCATION: Glyoxysomal.
CC -1- SIMILARITY: BELONGS TO THE ISOCITRATE LYASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z35499; CA84632.1; -
CC HSP: P28298; IDQU.
CC InterPro: IPR000918; Isocit_lyase.
CC Pfam: PF00463; ICL; 1.
CC ProDom: PD001857; Isocit_lyase; 2.
CC PROSITE: PS00161; ISOCITRATE_LYASE; 1.
CC PROSITE: PS00342; MICROBODIES_CTER; 1.
CC Glyoxylate bypass; Tricarboxylic acid cycle; Lyase; Glyoxysome.
CC ACT_SITE 213 213 PROBABLE.
CC FT SITE 574 576 MICROBODY TARGETING SIGNAL (POTENTIAL).
CC SQ SEQUENCE 576 AA; 64611 MW; 717D30B83CD81D6C CRC64;
CC -----
CC Query Match 41.6%; Score 44.5; DB 1; Length 576;
CC Best Local Similarity 30.8%; Pred. No. 9.7;
CC Matches 8; Conservative 10; Mismatches 3; Indels 5; Gaps 1;
CC -----
CC QY 1 DRSNLTSTSGHLSRR-----SDHLSR 21
CC Db 205 DQSSVTKKCGHMAGKVLVAVSEHNR 230
CC -----
CC RESULT 12
CC ACEA_GOSHI STANDARD; PRT; 576 AA.
CC ID ACEA_GOSHI
CC AC P17089;
CC DT 01-AUG-1990 (Rel. 15, Created)
CC DT 01-AUG-1990 (Rel. 15, Last sequence update)
CC DT 01-DEC-1992 (Rel. 24, Last annotation update)
CC DE Isocitrate lyase (EC 4.1.3.1) (Isocitrase) (Isocitratase) (ICL).
CC OS Gossypium hirsutum (Upland cotton).
CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
CC OC eurosids II; Malvales; Malvaceae; Gossypium.
CC OX NCBI_TaxID=3635;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=cv. Deltapine 62; TISSUE=Cotyledon;
CC RX MEDLINE=90304228; PubMed=2194576;
CC RA Turley R.B., Choe S.M., Trelease R.N.;
CC RT "Characterization of a cDNA clone encoding the complete amino acid
CC sequence of cotton isocitrate lyase."
CC RL Blochim. Biophys. Acta 1049:223-226(1990).
CC -1- FUNCTION: INVOLVED IN STORAGE LIPID MOBILIZATION DURING THE GROWTH
CC OF HIGHER PLANT SEEDLING.
CC -1- CATALYTIC ACTIVITY: Isocitrate - succinate + glyoxylate.
CC -1- PATHWAY: FIRST STEP IN GLYOXYLATE BYPASS, AN ALTERNATIVE TO THE
CC TRICARBOXYLIC ACID CYCLE (IN BACTERIA, PLANTS, AND FUNGI).
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Glyoxysomal.
CC -1- SIMILARITY: BELONGS TO THE ISOCITRATE LYASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X52136; CAA36381.1; -
CC FIRM: S10771; WZCNIU.
CC HSP: P28298; IDQU.
CC -----
CC Query Match 41.6%; Score 44.5; DB 1; Length 576;
CC Best Local Similarity 30.8%; Pred. No. 9.7;
CC Matches 8; Conservative 10; Mismatches 3; Indels 5; Gaps 1;
CC -----
CC QY 1 DRSNLTSTSGHLSRR-----SDHLSR 21
CC Db 205 DQSSVTKKCGHMAGKVLVAVSEHNR 230
CC -----
CC RESULT 13
CC ACEA_RICCO STANDARD; PRT; 576 AA.
CC ID ACEA_RICCO
CC AC P15479;
CC DT 01-APR-1990 (Rel. 14, Created)
CC DT 01-APR-1990 (Rel. 14, Last sequence update)
CC DT 01-DEC-1992 (Rel. 24, Last annotation update)
CC DE Isocitrate lyase (EC 4.1.3.1) (Isocitrase) (Isocitratase) (ICL).
CC OS Ricinus communis (Castor bean).
CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
CC OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
CC OX NCBI_TaxID=3988;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RA Beeching J.R., Northcote D.H.;
CC RT "Nucleic acid (cDNA) and amino acid sequences of isocitrate lyase from
CC castor bean."
CC RL Plant Mol. Biol. 8:471-475(1987).
CC -1- FUNCTION: INVOLVED IN STORAGE LIPID MOBILIZATION DURING THE GROWTH
CC OF HIGHER PLANT SEEDLING.
CC -1- CATALYTIC ACTIVITY: Isocitrate - succinate + glyoxylate.
CC -1- PATHWAY: FIRST STEP IN GLYOXYLATE BYPASS, AN ALTERNATIVE TO THE
CC TRICARBOXYLIC ACID CYCLE (IN BACTERIA, PLANTS, AND FUNGI).
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Glyoxysomal.
CC -1- SIMILARITY: BELONGS TO THE ISOCITRATE LYASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M17145; AAA53378.1; -
CC FIRM: S06274; WZCSI.
CC HSP: P28298; IDQU.
CC InterPro: IPR000918; Isocit_lyase.
CC Pfam: PF00463; ICL; 1.
CC ProDom: PD001857; Isocit_lyase; 2.
CC PROSITE: PS00161; ISOCITRATE_LYASE; 1.
CC PROSITE: PS00342; MICROBODIES_CTER; 1.
CC Glyoxylate bypass; Tricarboxylic acid cycle; Lyase; Glyoxysome.
CC ACT_SITE 213 213 PROBABLE.
CC FT SITE 574 576 MICROBODY TARGETING SIGNAL (POTENTIAL).
CC SQ SEQUENCE 576 AA; 64751 MW; 8177A7679050579B CRC64;
CC -----
CC Query Match 41.6%; Score 44.5; DB 1; Length 576;
CC Best Local Similarity 30.8%; Pred. No. 9.7;
CC Matches 8; Conservative 10; Mismatches 3; Indels 5; Gaps 1;
CC -----
CC QY 1 DRSNLTSTSGHLSRR-----SDHLSR 21
```

```
DR InterPro: IPR000918; Isocit_lyase.
DR Pfam: PF00463; ICL; 1.
DR ProDom: PD001857; Isocit_lyase; 2.
DR PROSITE: PS00161; ISOCITRATE_LYASE; 1.
DR PROSITE: PS00342; MICROBODIES_CTER; 1.
DR Glyoxylate bypass; Tricarboxylic acid cycle; Lyase; Glyoxysome.
DR ACT_SITE 213 213 PROBABLE.
DR FT SITE 574 576 MICROBODY TARGETING SIGNAL (POTENTIAL).
DR SQ SEQUENCE 576 AA; 64733 MW; 84D211AA3DCE8700 CRC64;
DR -----
DR Query Match 41.6%; Score 44.5; DB 1; Length 576;
DR Best Local Similarity 30.8%; Pred. No. 9.7;
DR Matches 8; Conservative 10; Mismatches 3; Indels 5; Gaps 1;
DR -----
DR QY 1 DRSNLTSTSGHLSRR-----SDHLSR 21
DR Db 205 DQSSVTKKCGHMAGKVLVAVSEHNR 230
DR -----
DR RESULT 13
DR ACEA_RICCO STANDARD; PRT; 576 AA.
DR ID ACEA_RICCO
DR AC P15479;
DR DT 01-APR-1990 (Rel. 14, Created)
DR DT 01-APR-1990 (Rel. 14, Last sequence update)
DR DT 01-DEC-1992 (Rel. 24, Last annotation update)
DR DE Isocitrate lyase (EC 4.1.3.1) (Isocitrase) (Isocitratase) (ICL).
DR OS Ricinus communis (Castor bean).
DR OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
DR OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
DR OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
DR OX NCBI_TaxID=3988;
DR RN [1]
DR RP SEQUENCE FROM N.A.
DR RA Beeching J.R., Northcote D.H.;
DR RT "Nucleic acid (cDNA) and amino acid sequences of isocitrate lyase from
DR castor bean."
DR RL Plant Mol. Biol. 8:471-475(1987).
DR -1- FUNCTION: INVOLVED IN STORAGE LIPID MOBILIZATION DURING THE GROWTH
DR OF HIGHER PLANT SEEDLING.
DR -1- CATALYTIC ACTIVITY: Isocitrate - succinate + glyoxylate.
DR -1- PATHWAY: FIRST STEP IN GLYOXYLATE BYPASS, AN ALTERNATIVE TO THE
DR TRICARBOXYLIC ACID CYCLE (IN BACTERIA, PLANTS, AND FUNGI).
DR -1- SUBUNIT: HOMOTETRAMER.
DR -1- SUBCELLULAR LOCATION: Glyoxysomal.
DR -1- SIMILARITY: BELONGS TO THE ISOCITRATE LYASE FAMILY.
DR -----
DR This SWISS-PROT entry is copyright. It is produced through a collaboration
DR between the Swiss Institute of Bioinformatics and the EMBL outstation -
DR the European Bioinformatics Institute. There are no restrictions on its
DR use by non-profit institutions as long as its content is in no way
DR modified and this statement is not removed. Usage by and for commercial
DR entities requires a license agreement (See http://www.isb-sib.ch/announce/
DR or send an email to license@isb-sib.ch).
DR -----
DR EMBL: M17145; AAA53378.1; -
DR FIRM: S06274; WZCSI.
DR HSP: P28298; IDQU.
DR InterPro: IPR000918; Isocit_lyase.
DR Pfam: PF00463; ICL; 1.
DR ProDom: PD001857; Isocit_lyase; 2.
DR PROSITE: PS00161; ISOCITRATE_LYASE; 1.
DR PROSITE: PS00342; MICROBODIES_CTER; 1.
DR Glyoxylate bypass; Tricarboxylic acid cycle; Lyase; Glyoxysome.
DR ACT_SITE 213 213 PROBABLE.
DR FT SITE 574 576 MICROBODY TARGETING SIGNAL (POTENTIAL).
DR SQ SEQUENCE 576 AA; 64751 MW; 8177A7679050579B CRC64;
DR -----
DR Query Match 41.6%; Score 44.5; DB 1; Length 576;
DR Best Local Similarity 30.8%; Pred. No. 9.7;
DR Matches 8; Conservative 10; Mismatches 3; Indels 5; Gaps 1;
DR -----
DR QY 1 DRSNLTSTSGHLSRR-----SDHLSR 21
```


Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	55	51.4	39	15	Q9EAB0	Q9eab0 simian t-ly
2	54	50.5	39	15	Q9QNl6	Q9qnl6 simian t-ly
3	54	50.5	87	15	Q80794	Q80794 human t-ly
4	54	50.5	87	15	Q9PX28	Q9px28 human t-ly
5	54	50.5	87	15	Q9PX16	Q9px16 human t-ly
6	54	50.5	240	15	Q9PX29	Q9px29 human t-ly
7	54	50.5	241	15	Q9PX17	Q9px17 human t-ly
8	53	49.5	532	5	Q9VC95	Q9vc95 drosophila
9	53	49.5	1765	5	Q8SX63	Q8sx63 drosophila
10	49	45.8	39	15	Q82210	Q82210 human t-ly
11	47	43.9	65	15	O12385	O12385 simian t-ly
12	47	43.9	636	3	Q9Y8A4	Q9y8a4 aspergillus
13	46	43.0	96	2	Q34547	Q34547 streptococc
14	46	43.0	205	4	Q9Y4U5	Q9y4u5 homo sapien
15	46	43.0	416	16	Q8XZC3	Q8xzC3 ralstonia s
16	46	43.0	996	5	Q95XN3	Q95xn3 caenorhabdi

```

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Tax protein (fragment).
GN TAX/REX.
OS Simian T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=33747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CH-REGINA;
RA Nipahs H., Verschuur E.J., Heeney J.L.;
RT "Reduced transmission and prevalence of HTLV-1/STLVcpz by captive
RT breeding of chimpanzees (Pan troglodytes).";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y18904; CAB52548.1; -.
FT NON_TER 1
FT TER 39
SQ SEQUENCE 39 AA; 4411 MW; 3F6A46266C885FEF CRC64;

Query Match 50.5%; Score 54; DB 15; Length 39;
Best Local Similarity 71.4%; Pred. No. 0.1;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 TTTSGHLSRRSDHL 19
Db :|:||||| ||||
24 SRPTGHLRSASDHL 37

RESULT 3
Q80794 PRELIMINARY; PRT; 87 AA.
AC Q80794;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE X protein.
GN X11.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88232270; PubMed=2897612;
RA Tsujimoto A., Teruuchi T., Inamura J., Shimotohno K., Miyoshi I.,
RA Miwa M.;
RT "Nucleotide sequence analysis of a provirus derived from HTLV-1-
RT associated myelopathy (HAM).";
RL Mol. Biol. Med. 5:29-42(1988).
DR EMBL; M37301; AAA45391.1; -.
SQ SEQUENCE 87 AA; 9963 MW; 141F3B7BA838B273 CRC64;

Query Match 50.5%; Score 54; DB 15; Length 87;
Best Local Similarity 71.4%; Pred. No. 0.23;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 TTTSGHLSRRSDHL 19
Db :|:||||| ||||
46 SRPTGHLRSASDHL 59

RESULT 4
Q9PXZ8 PRELIMINARY; PRT; 87 AA.
AC Q9PXZ8;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE P3 II protein.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93124536; PubMed=8419636;
RA Gessain A., Boeri E., Yanagihara R., Gallo R.C., Franchini G.;
RT "Complete nucleotide sequence of a highly divergent human T-cell
RT leukemia (lymphotropic) virus type I (HTLV-I) variant from melanesia:
RT genetic and phylogenetic relationship to HTLV-I strains from other
RT geographical regions.";
RL J. Virol. 67:1015-1023(1993).
SQ SEQUENCE 240 AA; 26561 MW; A17BD376692D2E37 CRC64;

Query Match 50.5%; Score 54; DB 15; Length 240;
Best Local Similarity 71.4%; Pred. No. 0.67;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

RA Gessain A., Boeri E., Yanagihara R., Gallo R.C., Franchini G.;
RT "Complete nucleotide sequence of a highly divergent human T-cell
RT leukemia (lymphotropic) virus type I (HTLV-I) variant from melanesia:
RT genetic and phylogenetic relationship to HTLV-I strains from other
RT geographical regions.";
RL J. Virol. 67:1015-1023(1993).
SQ SEQUENCE 87 AA; 9988 MW; 34A239AC7EE7D9CD CRC64;

Query Match 50.5%; Score 54; DB 15; Length 87;
Best Local Similarity 71.4%; Pred. No. 0.23;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 TTTSGHLSRRSDHL 19
Db :|:||||| ||||
46 SRPTGHLRSASDHL 59

RESULT 5
Q9PXY6 PRELIMINARY; PRT; 87 AA.
AC Q9PXY6;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE PX-ORF II fusion protein.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92409607; PubMed=1528897;
RA Korainik I.J., Gessain A., Klotman M.E., Lo Monaco A., Berneman Z.N.,
RA Franchini G.;
RT "Protein isoforms encoded by the px region of human T-cell
RT leukemia/lymphotropic virus type I.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:8813-8817(1992).
SQ SEQUENCE 87 AA; 10080 MW; 12DDEBDB6838B277 CRC64;

Query Match 50.5%; Score 54; DB 15; Length 87;
Best Local Similarity 71.4%; Pred. No. 0.23;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 TTTSGHLSRRSDHL 19
Db :|:||||| ||||
46 SRPTGHLRSASDHL 59

RESULT 6
Q9PXZ9 PRELIMINARY; PRT; 240 AA.
AC Q9PXZ9;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE P30 II protein.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93124536; PubMed=8419636;
RA Gessain A., Boeri E., Yanagihara R., Gallo R.C., Franchini G.;
RT "Complete nucleotide sequence of a highly divergent human T-cell
RT leukemia (lymphotropic) virus type I (HTLV-I) variant from melanesia:
RT genetic and phylogenetic relationship to HTLV-I strains from other
RT geographical regions.";
RL J. Virol. 67:1015-1023(1993).
SQ SEQUENCE 240 AA; 26561 MW; A17BD376692D2E37 CRC64;

Query Match 50.5%; Score 54; DB 15; Length 240;
Best Local Similarity 71.4%; Pred. No. 0.67;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      6 TRTSGHLSRRSDHL 19
Db      199 SRPTGHLRSRSDHL 212

RESULT 7
Q9PXV7 ID Q9PXV7 PRELIMINARY; PRT; 241 AA.
AC Q9PXV7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE PX-TAX-ORF II fusion protein.
OS Human T-lymphotropic virus 1.
OC Viruses; Retrovirdae; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92409607; PubMed=1528897;
RA Koralnik I.J., Gessain A., Klotman M.E., Lo Monaco A., Berneman Z.N.,
RA Franchini G.;
RT "Protein isoforms encoded by the pX region of human T-cell
RT leukemia/lymphotropic virus type 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:8813-8817(1992).
SQ SEQUENCE 241 AA; 26791 MW; 7607F1F66C55CF5E CRC64;

Query Match 50.5%; Score 54; DB 15; Length 241;
Best Local Similarity 71.4%; Pred. No. 0.67;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      6 TRTSGHLSRRSDHL 19
Db      200 SRPTGHLRSRSDHL 213

RESULT 8
Q9VC95 ID Q9VC95 PRELIMINARY; PRT; 532 AA.
AC Q9VC95;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE CG18670 protein.
GN CG6400 OR CG18670.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

```

```

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Kobayashi C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Fan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003747; AAF56279.1; -.
DR HSP; Q92831; 1891.
DR FlyBase; FBgn0039181; CG6400.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; Bromodomain; 1.
DR PRINTS; PR00503; BROMODOMAIN.
DR PRINTS; PR01574; TUBBYPROTEIN.
DR SMART; SM00297; BROMO; 1.
DR PROSITE; PS00114; BROMODOMAIN_2; 1.
SQ SEQUENCE 532 AA; 58615 MW; F2A7D2861E1F3469 CRC64;

Query Match 49.5%; Score 53; DB 5; Length 532;
Best Local Similarity 58.8%; Pred. No. 2.2;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy      2 RSNLTRSGHLSRRSDH 18
Db      380 RANGTRTSSHLGAENDH 396

RESULT 9
Q8SX63 ID Q8SX63 PRELIMINARY; PRT; 1765 AA.
AC Q8SX63;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE LD40380p.
GN CG6400.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY094828; AAM11181.1; -.
SQ SEQUENCE 1765 AA; 195942 MW; 6AF1A6933FA482C3 CRC64;

Query Match 49.5%; Score 53; DB 5; Length 1765;
Best Local Similarity 58.8%; Pred. No. 7.7;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy      2 RSNLTRSGHLSRRSDH 18
Db      1299 RANGTRTSSHLGAENDH 1315

```

```

RESULT 10
Q82210 PRELIMINARY; PRT; 39 AA.
AC Q82210;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Tax protein (Fragment).
GN TAX.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSHR-1;
RA Yanagihara R., Saitou N., Nerurkar V.R., Song K.J., Bastian I.,
RA Franchini G., Gajdusek D.C.;
RT "Molecular phylogeny and dissemination of human T-cell lymphotropic
RT virus type I viewed within the context of primate evolution and human
RT migration.";
RL Cell. Mol. Biol. 0:0-0(1995).
DR EMBL; U12121; AAA99642.1; -.
FT NON_TER 1 1
FT 39 39
SQ SEQUENCE 39 AA; 4425 MW; 23A6CB4626765F9B CRC64;

Query Match 45.88; Score 49; DB 15; Length 39;
Best Local Similarity 64.38; Pred. No. 0.87;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 TRTSGHLSRRSDHL 19
: | | | | | | |
Db 23 SRPTGHLPTSDDL 36

RESULT 11
Q12385 PRELIMINARY; PRT; 65 AA.
AC Q12385;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Rex p13/p30 protein (Fragment).
GN REX.
OS Simian T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=33747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KA;
RX MEDLINE=98120820; PubMed=9460922;
RA Verschoor E.J., Warren K., Niphuis H., Heriyanto, Swan R.A.,
RA Heeney J.L.;
RT "Characterization of a simian T-lymphotropic virus from a wild-caught
RT orangutan (Pongo pygmaeus) from Kalimantan, Indonesia.";
RL J. Gen. Virol. 79:51-55(1998).
DR EMBL; Y13146; CAA73607.1; -.
FT NON_TER 1 1
FT 65 65
SQ SEQUENCE 65 AA; 7434 MW; 044211FEBF74AA65 CRC64;

Query Match 43.98; Score 47; DB 15; Length 65;
Best Local Similarity 64.38; Pred. No. 2.4;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 TRTSGHLSRRSDHL 19
: | | | | | | |
Db 24 SRPTGHLRASGHL 37

RESULT 12
Q9Y8A4 PRELIMINARY; PRT; 636 AA.
ID Q9Y8A4
AC Q9Y8A4;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE 5-aminolevulinic acid synthase.
GN HEMA.
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A1560;
RA Elrod S.L., Jones A., Cherry J.R.;
RT "Cloning of 5-aminolevulinic synthase from Aspergillus oryzae and its
RT use as a selectable marker.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152374; AAD38391.1; -.
DR InterPro; IPR003408; Ala_synthase.
DR InterPro; IPR004839; AminoTransf1/2.
DR InterPro; IPR001917; NHtransf.2.
DR Pfam; PF02490; Ala_synthase; 1.
DR Pfam; PF00155; aminotran_1.2; 1.
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; 1.
SQ SEQUENCE 636 AA; 68145 MW; F70BC2A073A0CDBA CRC64;

Query Match 43.98; Score 47; DB 3; Length 636;
Best Local Similarity 42.18; Pred. No. 26;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 DRSNLTRTSGHLSRRSDHL 19
: | | | | | | |
Db 531 ERLRITPTGHIKEHRDHL 549

RESULT 13
Q54547 PRELIMINARY; PRT; 96 AA.
ID Q54547
AC Q54547;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Emml protein (Fragment).
GN EMMML.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M TYPE 30;
RA Whatmore A.M.;
RT "Sequence Analysis of the Emm-Like Gene Family of Streptococcus
RT Pyogenes.";
RL Thesis (1993), Microbiology, University of Newcastle Upon Tyne.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=M TYPE 30;
RX MEDLINE=95198537; PubMed=7891551;
RA Whatmore A.M., Kapur V., Sullivan D.J., Musser J.M., Kehoe M.A.;
RT "Non-congruent relationships between variation in emm gene sequences
RT and the population genetic structure of group A streptococci.";
RL Mol. Microbiol. 14:619-631(1994).
DR EMBL; U11944; AAA99560.1; -.
FT NON_TER 1 1
FT 96 96
SQ SEQUENCE 96 AA; 10640 MW; E6A9BC38FFB4B865 CRC64;

Query Match 43.08; Score 46; DB 2; Length 96;
Best Local Similarity 38.18; Pred. No. 5.3;
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 DRSNLTRTSGHLSRRSDHL 21

```


Query Match	43.0%	Score 46;	DB 16;	Length 416;
Best Local Similarity	53.3%	Pred. No. 24;		
Matches 8;	Conservative	3;	Mismatches	4;
			Indels	0;
			Gaps	0;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 16:30:01 : Search time 26 Seconds
(without alignments)
23.765 Million cell updates/sec

Title: SEQ55-247-68

Perfect score: 107

Sequence: 1 DRSNLRTSGHLRRSDHLR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*
1: /cgn2.6/ptodata/1/1aa/5A_COMB.pap.*
2: /cgn2.6/ptodata/1/1aa/5B_COMB.pap.*
3: /cgn2.6/ptodata/1/1aa/6A_COMB.pap.*
4: /cgn2.6/ptodata/1/1aa/6B_COMB.pap.*
5: /cgn2.6/ptodata/1/1aa/PCIRUS_COMB.pap.*
6: /cgn2.6/ptodata/1/1aa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	54.2	21	4 US-09-229-007A-97	Sequence 97, Appl
2	54	50.5	40	1 US-07-743-518-18	Sequence 18, Appl
3	53	49.5	21	4 US-09-229-007A-94	Sequence 94, Appl
4	51	47.7	21	4 US-09-229-007A-96	Sequence 96, Appl
5	47	43.9	636	2 US-08-871-266B-2	Sequence 2, Appl
6	47	43.9	636	2 US-08-819-458A-2	Sequence 2, Appl
7	47	43.9	636	2 US-09-018-864A-2	Sequence 2, Appl
8	47	43.9	636	3 US-08-871-267B-2	Sequence 2, Appl
9	47	43.9	636	4 US-09-618-419-2	Sequence 2, Appl
10	46	43.0	21	4 US-09-229-007A-95	Sequence 95, Appl
11	42	39.3	169	4 US-08-856-841-11	Sequence 11, Appl
12	42	39.3	649	2 US-08-871-266B-16	Sequence 16, Appl
13	42	39.3	649	2 US-09-018-864A-16	Sequence 16, Appl
14	42	39.3	649	3 US-08-871-267B-22	Sequence 22, Appl
15	42	39.3	649	4 US-09-618-419-22	Sequence 22, Appl
16	41	38.3	312	4 US-08-961-083-12	Sequence 12, Appl
17	41	38.3	478	3 US-09-147-522-2	Sequence 2, Appl
18	40	37.4	21	4 US-09-229-007A-93	Sequence 93, Appl
19	40	37.4	484	4 US-09-134-001C-3116	Sequence 3116, Ap
20	40	37.4	1050	3 US-09-045-632-49	Sequence 49, Appl
21	40	37.4	1050	3 US-09-045-632-50	Sequence 50, Appl
22	40	37.4	1833	4 US-08-621-944A-4	Sequence 4, Appl
23	40	37.4	1833	4 US-08-945-567D-4	Sequence 4, Appl
24	40	37.4	1992	4 US-08-621-944A-3	Sequence 3, Appl
25	40	37.4	1992	4 US-08-945-567D-3	Sequence 3, Appl
26	40	37.4	2048	4 US-09-268-347-48	Sequence 48, Appl
27	39	36.4	218	4 US-09-068-655-7	Sequence 7, Appl

28 39 36.4 376 1 US-08-303-238-1 Sequence 1, Appl
29 39 36.4 376 4 US-08-458-834-1 Sequence 1, Appl
30 39 36.4 783 1 US-08-843-521-2 Sequence 2, Appl
31 39 36.4 783 4 US-09-012-871-2 Sequence 2, Appl
32 38.5 36.0 560 2 US-08-981-690-2 Sequence 2, Appl
33 38 35.5 634 1 US-07-872-644-51 Sequence 51, Appl
34 38 35.5 634 1 US-08-297-494-51 Sequence 51, Appl
35 38 35.5 634 1 US-08-297-510-51 Sequence 51, Appl
36 38 35.5 634 1 US-08-479-532-51 Sequence 51, Appl
37 38 35.5 634 1 US-08-455-526-51 Sequence 51, Appl
38 38 35.5 634 1 US-08-455-525-51 Sequence 51, Appl
39 38 35.5 634 3 US-09-139-491-51 Sequence 51, Appl
40 38 35.5 634 5 PCT-U892-03222-51 Sequence 51, Appl
41 38 35.5 725 1 US-08-448-170-4 Sequence 4, Appl
42 38 35.5 725 4 US-08-961-803-7 Sequence 7, Appl
43 37.5 35.0 907 3 US-08-938-830-26 Sequence 26, Appl
44 37.5 35.0 907 3 US-09-020-222-26 Sequence 26, Appl
45 37 34.6 236 4 US-08-711-417C-175 Sequence 175, App

ALIGNMENTS

RESULT 1

US-09-229-007A-97

; Sequence 97, Application US/09229007A

; Patent No. 6453242

; GENERAL INFORMATION:

; APPLICANT: Eisenberg, Stephen P.

; APPLICANT: Case, Casey C.

; APPLICANT: Cox III, George N.

; APPLICANT: Jamieson, Andrew

; APPLICANT: Rebar, Edward J.

; APPLICANT: Sangamo Biosciences, Inc.

; TITLE OF INVENTION: Selection of Sites for Targeting by Zinc Finger

; TITLE OF INVENTION: Proteins and Methods of Designing Zinc Finger Proteins

; FILE REFERENCE: 019496-001800US

; CURRENT APPLICATION NUMBER: US/09/229,007A

; NUMBER OF SEQ ID NOS: 97

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 97

; LENGTH: 21

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:ZFP sequence

; OTHER INFORMATION: (F1, F2, F3) from SBS design GL-8.3.1

US-09-229-007A-97

Query Match 54.2%; Score 58; DB 4; Length 21;

Best Local Similarity 55.0%; Pred. No. 0.0025;

Matches 11; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 RSNLRTSGHLRRSDHLR 21

Db 2 KDSLVRTSDHLASRSDNLR 21

RESULT 2

US-07-743-518-18

; Sequence 18, Application US/07743518

; Patent No. 5397696

; GENERAL INFORMATION:

; APPLICANT: YANAGIHARA, RICHARD

; APPLICANT: NERURKAR, VIVEK R.

; APPLICANT: JENKINS, CAROL

; APPLICANT: MILLER, MARK

; APPLICANT: GARRUTO, RALPH M.

; TITLE OF INVENTION: PAPUA NEW GUINEA HUMAN T-LYMPHOTROPIC

; TITLE OF INVENTION: VIRUS

; NUMBER OF SEQUENCES: 26

TELEPHONE: 212-878-9652
TELEFAX: 212-878-9655
TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 636 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-871-266B-2

Query Match 43.9%; Score 47; DB 2; Length 636;

Best Local Similarity 42.1%; Pred. No. 8.3;

Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 DRSNLTSGHLSRRSDHL 19

:|:|:|:|

Db 531 ERLRITPTPGHIKEHRDHL 549

RESULT 6

US-08-819-458A-2

; Sequence 2, Application US/08819458A

; Patent No. 5891669

; GENERAL INFORMATION:

; APPLICANT: Jensen, Ejner B.

; APPLICANT: Cherry, Joel

; APPLICANT: Elrod, Susan L.

; TITLE OF INVENTION: Methods For Producing Polypeptides

; TITLE OF INVENTION: In Respiratory-Deficient Cells

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 5891669 No. 5891669disk Of No. 5891669th America, Inc.

; STREET: 405 Lexington Avenue

; CITY: New York

; STATE: NY

; COUNTRY: U.S.A.

; ZIP: 10174

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/819,458A

; FILING DATE: 17-MAR-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Lambiris, Elias J.

; REGISTRATION NUMBER: 33,728

; REFERENCE/DOCKET NUMBER: 5215.000-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-867-0123

; TELEFAX: 212-878-9655

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 636 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-819-458A-2

Query Match

Best Local Similarity 43.9%; Score 47; DB 2; Length 636;

Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 DRSNLTSGHLSRRSDHL 19

:|:|:|:|

Db 531 ERLRITPTPGHIKEHRDHL 549

RESULT 7

US-09-018-864A-2

; Sequence 2, Application US/09018864A

; Patent No. 5958747

; GENERAL INFORMATION:

; APPLICANT: Elrod, Susan L.

; APPLICANT: Cherry, Joel R.

; TITLE OF INVENTION: Aspergillus Oryzae 5-Aminolevulinic

; TITLE OF INVENTION: Acid Synthases And Nucleic Acids Encoding Same

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 5958747 No. 5958747disk Of No. 5958747th America, Inc.

; STREET: 405 Lexington Avenue - 64th Fl.

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10174

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/018,864A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/871,266

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Lambiris, Elias J.

; REGISTRATION NUMBER: 33,728

; REFERENCE/DOCKET NUMBER:

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-878-9652

; TELEFAX: 212-878-9655

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 636 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

US-09-018-864A-2

Query Match

Best Local Similarity 43.9%; Score 47; DB 2; Length 636;

Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 DRSNLTSGHLSRRSDHL 19

:|:|:|:|

Db 531 ERLRITPTPGHIKEHRDHL 549

RESULT 8

US-08-371-267B-2

; Sequence 2, Application US/08871267B

; Patent No. 6100057

; GENERAL INFORMATION:

; APPLICANT: Elrod, Susan L.

; APPLICANT: Cherry, Joel R.

; TITLE OF INVENTION: A Method for Increasing Hemoprotein

; TITLE OF INVENTION: Production in Filamentous Fungi

; NUMBER OF SEQUENCES: 39

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6100057 No. 6100057disk Of No. 6100057th America, Inc.

; STREET: 405 Lexington Avenue - 64th Fl.

; CITY: New York

CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 169
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE: 1
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 11:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: FRITSCH, E.F.
AUTHORS: SAMBROOK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:

PAGES: 1982
DATE: 1982
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 11:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
JOURNAL: GENE
VOLUME:
ISSUE: 67
PAGES: 31 - 40
DATE: 1988
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 11:
US-08-856-841-11
Query Match 39.3%; Score 42; DB 4; Length 169;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 10 GHLRRSDHLSR 21
| | | | | | | | | |
Db 1 GILSRPDHMSK 12
RESULT 12
US-08-871-266B-16
; Sequence 16, Application US/08671266B
; Patent No. 5871991
; GENERAL INFORMATION:
; APPLICANT: Elrod, Susan L.
; TITLE OF INVENTION: Aspergillus Oryzae 5-Aminolevulinic
; TITLE OF INVENTION: Acid Synthases And Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSER: No. 5871991o No. 5871991disk Of No. 5871991th America, Inc.
; STREET: 405 Lexington Avenue - 64th Fl.
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,266B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-878-9652
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5871991e

US-08-871-266B-16

Query Match 39.3%; Score 42; DB 2; Length 649;
Best Local Similarity 42.1%; Pred. No. 56;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

```
QY      1 DRSNLTRTSGHLSRRSDHL 19
        :| :| :| :| :|
Db     533 ERLRITPTPGHTQELRDHL 551
```

RESULT 13

US-09-018-864A-16
Sequence 16, Application US/09018864A
Patent No. 5958747
GENERAL INFORMATION:
APPLICANT: Elrod, Susan L.
APPLICANT: Cherry, Joel R.
TITLE OF INVENTION: Aspergillus Oryzae 5-Aminolevulinic
Acid Synthases And Nucleic Acids Encoding Same
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958747o No. 5958747disk Of No. 5958747th America, Inc.
STREET: 405 Lexington Avenue - 64th Fl.

Query Match	39.3%	Score 42;	DB 2;	Length 649;
Best Local Similarity	42.1%	Pred. NO. 56;		
Matches 8;	Conservative	2;	Mismatches	9;
			Indels	0;
			Gaps	0;

```

QY      1 DRSNLTRTSGHLSRRSDHL 19
      :|:|:|:|
db     533 ERLRITPTPGHTOELRDHL 551

```

RESULT 14

US-08-871-267B-22
; Sequence 22, Application US/08871267B
; Patent No. 610057
; GENERAL INFORMATION:
; APPLICANT: Elrod, Susan L.
; APPLICANT: Cherry, Joel R.

APPLICANT: Jones, Aubrey
 TITLE OF INVENTION: A Method for Increasing Hemoprotein
 Production in Filamentous Fungi
 NUMBER OF SEQUENCES: 39
 CORRESPONDENCE ADDRESS:
 ADDRESS: No. 61000570 No. 61000570disk Of No. 61000570 America, Inc.
 STREET: 405 Lexington Avenue - 64th Fl.
 CITY: New York, N.Y. 10017-2498
 COUNTRY: U.S.A.

Query Match 39.3%; Score 42; DB 3; Length 649;
Best Local Similarity 42.1%; Pred. No. 56;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

```
Qy      1 DRSNLTRTSGHLSRRSDHL 19
        : | : | | | | |
Db     533 ERLRITTPGHTOELRDHL 551
```

RESULT 15

```

RESOLUTION 1
US-09-618-419-22
; Sequence 22, Application US/09618419
; Patent No. 6261827
; GENERAL INFORMATION:
; APPLICANT: Elrod, Susan L.
; Cherry, Joel R.
; Jones, Aubrey
; TITLE OF INVENTION: A Method for Increasing Hemoprotein
; Production in Filamentous Fungi
;

```

COMPUTER RECEIPT FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for W1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/O
FILING DATE: 18-Jul-2000
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/871,267
FILING DATE: 9-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Rozek, Carol E.
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 4771.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-878-9652
TELEFAX: 212-878-9655
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 649 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6261827e
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-618-419-22

Query Match 39.3%; Score 42; DB 4; Length 649;
Best Local Similarity 42.1%; Pred. No. 56;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 DRNLRTTSGHLRRSDHL 19
: : : : :
Db 533 ERLRITPTPGHTQLRDHL 551

Search completed: July 15, 2003, 16:31:02
Job time : 27 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2003, 16:30:36 ; Search time 21 Seconds
(without alignments)
116.420 Million cell updates/sec

Title: SEQ55-247-68

Perfect score: 107

Sequence: 1 DRSNLTSTGHLRRSDHLSR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	54.2	21	9	US-10-113-424-97
2	53	49.5	21	9	Sequence 97, Appl
3	51	47.7	21	9	Sequence 94, Appl
4	46	43.0	21	9	Sequence 96, Appl
5	45	42.1	1610	9	Sequence 95, Appl
6	44.5	41.6	451	9	Sequence 6666, Ap
7	43	40.2	105	10	Sequence 5, Appli
8	42	39.3	636	9	Sequence 11688, A
9	41	38.3	140	9	Sequence 7810, Ap
10	41	38.3	140	10	Sequence 635, Appl
11	41	38.3	140	10	Sequence 635, Appl
12	41	38.3	312	10	Sequence 635, App
13	41	38.3	335	9	Sequence 12, Appl
14	41	38.3	552	9	Sequence 4, Appli
15	41	38.3	552	9	Sequence 196, App
16	41	38.3	552	9	Sequence 196, App
17	41	38.3	552	9	Sequence 196, App
18	41	38.3	552	9	Sequence 196, App
19	41	38.3	552	9	Sequence 196, App

20	41	38.3	552	9	US-10-176-921-196	Sequence 196, App
21	41	38.3	552	9	US-10-137-865-196	Sequence 196, App
22	41	38.3	552	9	US-10-140-474-196	Sequence 196, App
23	41	38.3	552	9	US-10-142-431-196	Sequence 196, App
24	41	38.3	552	9	US-10-143-114-196	Sequence 196, App
25	41	38.3	552	9	US-10-140-002-196	Sequence 196, App
26	41	38.3	552	9	US-10-142-419-196	Sequence 196, App
27	41	38.3	552	9	US-10-123-262-196	Sequence 196, App
28	41	38.3	552	9	US-10-142-423-196	Sequence 196, App
29	41	38.3	552	9	US-10-121-050-196	Sequence 196, App
30	41	38.3	552	9	US-10-141-755-196	Sequence 196, App
31	41	38.3	552	9	US-10-143-032-196	Sequence 196, App
32	41	38.3	552	9	US-10-123-108-196	Sequence 196, App
33	41	38.3	552	9	US-10-123-236-196	Sequence 196, App
34	41	38.3	552	9	US-10-123-261-196	Sequence 196, App
35	41	38.3	552	9	US-10-140-921-196	Sequence 196, App
36	41	38.3	552	9	US-10-140-928-196	Sequence 196, App
37	41	38.3	552	9	US-10-121-045-196	Sequence 196, App
38	41	38.3	552	9	US-10-123-292-196	Sequence 196, App
39	41	38.3	552	9	US-10-123-903-196	Sequence 196, App
40	41	38.3	552	9	US-10-124-819-196	Sequence 196, App
41	41	38.3	552	9	US-10-124-822-196	Sequence 196, App
42	41	38.3	552	9	US-10-140-925-196	Sequence 196, App
43	41	38.3	552	9	US-10-160-498-196	Sequence 196, App
44	41	38.3	552	9	US-10-121-041-196	Sequence 196, App
45	41	38.3	552	9	US-10-121-043-196	Sequence 196, App

ALIGNMENTS

RESULT 1
US-10-113-424-97
; Sequence 97, Application US/10113424
; Publication No. US20030105593A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Case, Casey C.
; APPLICANT: Cox III, George N.
; APPLICANT: Jamieson, Andrew
; APPLICANT: Rebar, Edward J.
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Selection of Sites for Targeting by Zinc Finger
; TITLE OF INVENTION: Proteins and Methods of Designing Zinc Finger Proteins
; TITLE OF INVENTION: To Bind to Preslected Sites
; FILE REFERENCE: 019496-00180005
; CURRENT APPLICATION NUMBER: US/10/113,424
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US/09/229,007A
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 97
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:ZFP sequence
; OTHER INFORMATION: (F1, F2, F3) from SBS design GL-8.3.1
US-10-113-424-97
Query Match 54.2%; Score 58; DB 9; Length 21;
Best Local Similarity 55.0%; Pred. No. 0.0088;
Matches 11; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 2 RSNLTSTGHLRRSDHLSR 21
: : ||| ||: |||: |
Db 2 KDSLVRTSDHLASRSDNLTR 21
RESULT 2
US-10-113-424-94
; Sequence 94, Application US/10113424

Query Match 42.1%; Score 45; DB 9; Length 1610;
Best Local Similarity 42.9%; Pred. No. 1.4e+02;
Matches 9; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 DRSNLTSGHLSRRSDHLSR 21
DB 1588 DEPEIGTVGAHLSRRIDEISR 1608

RESULT 6

US-10-071-894-5
; Sequence 5, Application US/10071894
; Publication No. US20030082669A1
; GENERAL INFORMATION:
; APPLICANT: Lorenz, Michael C.
; APPLICANT: Fink, Gerald R.
; TITLE OF INVENTION: Glyoxylate Cycle Enzymes as Targets for
; TITLE OF INVENTION: Antifungal Drug Development
; FILE REFERENCE: 0399.2026-001
; CURRENT APPLICATION NUMBER: US/10/071,894
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,622
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 451
; TYPE: PRT
; ORGANISM: A. thaliana Acea
US-10-071-894-5

Query Match 41.6%; Score 44.5; DB 9; Length 451;
Best Local Similarity 30.8%; Pred. No. 40;
Matches 8; Conservative 10; Mismatches 3; Indels 5; Gaps 1;

QY 1 DRSNLTSGHLSRR-----SDHLSR 21
DB 202 DQSSVTKKGGMAGKVLVAVSEHINR 227

RESULT 7

US-09-815-242-11688
; Sequence 11688, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11688
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-815-242-11688

Query Match 40.2%; Score 43; DB 10; Length 105;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 4 NLRTSGHLSRRSDHLSR 21
DB 49 NLVRSSGRVSFLNDDISR 56

RESULT 8

US-10-156-761-7810
; Sequence 7810, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7810
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-7810

Query Match 39.3%; Score 42; DB 9; Length 636;
Best Local Similarity 46.7%; Pred. No. 1.5e+02;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 5 LTRTSGHLSRRSDHL 19
DB 475 LTRINGNATRNSHI 489

RESULT 9

US-10-091-504-635
; Sequence 635, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; PRIOR APPLICATION removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 635
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-091-504-635

Query Match 38.3%; Score 41; DB 9; Length 140;

```
Best Local Similarity 64.3%; Pred. No. 38;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 SNLRTSGHLSRRS 16
Db 110 SNVTSTSGHSGSS 123

RESULT 10
US-09-908-711-69
; Sequence 69, Application US/09908711
; Patent No. US20020045230A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; CURRENT APPLICATION NUMBER: US/09/908,711
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US01/01360
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,867
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01344
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01345
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,888
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01329
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,905
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01354
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,891
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01339
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01340
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,874
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01334
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,898
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01320
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,853
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01349
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,902
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01239
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,870
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01348
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,882
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01347
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,896
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01307
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,864

; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01341
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,856
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01336
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,868
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01312
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-908-711-69

Query Match 38.3%; Score 41; DB 10; Length 140;
Best Local Similarity 64.3%; Pred. No. 38;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 SNLRTSGHLSRRS 16
Db 110 SNVTSTSGHSGSS 123

RESULT 11
US-09-764-869-635
; Sequence 635, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 635
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-869-635

Query Match 38.3%; Score 41; DB 10; Length 140;
Best Local Similarity 64.3%; Pred. No. 38;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 SNLRTSGHLSRRS 16
Db 110 SNVTSTSGHSGSS 123

RESULT 12
US-09-765-272-12
; Sequence 12, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
```

CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 312 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-765-272-12

Query Match 38.3%; Score 41; DB 10; Length 312;
Best Local Similarity 61.5%; Pred. No. 95;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 8 TSGHLSRRSDHLS 20
|||:||||:|
DB 98 TSGIISRKSDNVS 110

RESULT 13
US-09-769-787-4
Sequence 4, Application US/09769787
Publication No. US20030091577A1
GENERAL INFORMATION:
APPLICANT: Microbial Technics Limited
APPLICANT: Gilbert, Christophe FG
APPLICANT: Hansbro, Philip M
TITLE OF INVENTION: Proteins
FILE REFERENCE: PWC/P21129WO
CURRENT APPLICATION NUMBER: US/09/769,787
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: GB 9816337.1
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: US 60/125164
PRIOR FILING DATE: 1998-03-19
NUMBER OF SEQ ID NOS: 388
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 335
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-769-787-4

Query Match 38.3%; Score 41; DB 9; Length 335;
Best Local Similarity 61.5%; Pred. No. 1e+02;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 8 TSGHLSRRSDHLS 20
|||:||||:|
DB 121 TSGIISRKSDNVS 133

RESULT 14
US-10-028-072-196
Sequence 196, Application US/10028072
Publication No. US20030004311A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang
TITLE OF INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028,072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561

; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081695
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081818
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082999
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Query Match 38.3%; Score 41; DB 9; Length 552;
Best Local Similarity 36.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 9; Indels 0; Caps 0;

QY 1 DRSNLTRTSCHLSRSDHL 19
|| | | | |
Db 498 DROQWTKGSHIEHIAHL 516

RESULT 15
US-10-121-049-196
; Sequence 196, Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Flivaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 196
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-049-196

Query Match 38.3%; Score 41; DB 9; Length 552;
Best Local Similarity 36.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Oy 1 DRSNLRTSGHLSRSDHL 19
||| | | | | | | | | |
Db 498 DROQWTKGSHIEHIAHL 516

Search completed: July 15, 2003, 16:34:36
Job time : 23 secs

THIS PAGE BLANK (USPTO)